(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 July 2001 (26.07.2001)

PCT

(10) International Publication Number
WO 01/53312 A1

(51) International Patent Classification[7]: C07H 21/04, C12N 15/11, 15/63, 15/70, 15/82, 15/85, C07K 14/00

(21) International Application Number: PCT/US00/34263

(22) International Filing Date:
26 December 2000 (26.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/488,725	21 January 2000 (21.01.2000)	US
09/552,317	25 April 2000 (25.04.2000)	US
09/598,042	9 July 2000 (09.07.2000)	US
09/620,312	19 July 2000 (19.07.2000)	US
09/653,450	3 August 2000 (03.08.2000)	US
09/662,191	14 September 2000 (14.09.2000)	US
09/693,036	19 October 2000 (19.10.2000)	US
09/727,344	29 November 2000 (29.11.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/488,725 (CIP)
Filed on	21 January 2000 (21.01.2000)
US	09/552,317 (CIP)
Filed on	25 April 2000 (25.04.2000)
US	09/598,042 (CIP)
Filed on	9 July 2000 (09.07.2000)
US	09/620,312 (CIP)
Filed on	19 July 2000 (19.07.2000)
US	09/653,450 (CIP)
Filed on	3 August 2000 (03.08.2000)
US	09/662,191 (CIP)
Filed on	14 September 2000 (14.09.2000)
US	09/693,036 (CIP)
Filed on	19 October 2000 (19.10.2000)
US	09/727,344 (CIP)
Filed on	29 November 2000 (29.11.2000)

(71) Applicant *(for all designated States except US)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **LIU, Chenghua** [CN/US]; 1125 Ranchero Way #14, San Jose, CA 95117 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **CHEN, Rui-hong** [US/US]; 1031 Flying Fish Street, Foster City, CA 94404 (US). **MA, Yunqing** [CN/US]; 280 W. California Avenue #206, Sunnyvale, CA 94086 (US). **QIAN, Xiaohong, B.** [CN/US]; 3662 Tumble Way, San Jose, CA 95132 (US). **REN, Feiyan** [US/US]; 7703 Oak Meadow Court, Cupertino, CA 95014 (US). **WANG, Dunrui** [CN/US]; 932 La Palma, Milpitas, CA 95035 (US). **WANG, Jian-Rui** [CN/US]; 744 Stendhal Lane, Cupertino, CA 95014 (US). **WANG, Zhiwei** [CN/US]; 836 Alturas Avenue, B36, Sunnyvale, CA 94085 (US). **WEHRMAN, Tom** [US/US]; 3210 CCSR Mol Pharm, 269 W. Campus Drive, Stanford, CA 94305 (US). **XU, Chongjun** [CN/US]; 4918 Manitoba Drive, San Jose, CA 95130 (US). **XUE, Aidong, J.** [CN/US]; 1621 S. Mary Avenue, Sunnyvale, CA 94087 (US). **YANG, Yonghong** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **ZHAO, Qing, A.** [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). **ZHOU, Ping** [CN/US]; 1461 Japaul Lane, San Jose, CA 95132 (US). **GOODRICH, Ryle** [US/US]; 4896 Sandy Lane, San Jose, CA 95124 (US). **DRMANAC, Radoje, T.** [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
with international search report
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

[Continued on next page]

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

WO 01/53312 A1

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1786 and 3573-5358. The polypeptides sequences are designated SEQ ID NO: 2n (wherein n = 1 to 20). The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1786 and 3573-5358 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1786 and 3573-5358. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1786 and 3573-5358 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1786 and 3573-5358. The sequence information can be a segment of any one of SEQ ID NO:1-1786 and 3573-5358 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1786 and 3573-5358.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1786 and 3573-5358 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1786 and 3573-5358 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1786 and 3573-5358; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1-1786 and 3573-5358; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO:1-1786 and 3573-5358. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1786 and 3573-5358; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1786 and 3573-5358; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g., in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2); for which they have a signature region (as set forth in Table 3); or for which they have homology to a gene family (as set forth in Table 4). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady

and continuous source of germ cells for the production of gametes. The term "primordial germ cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30

nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-20.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1786 and 3573-5358. The sequence information can be a segment of any one of SEQ ID NO:1-1786 and 3573-5358 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1786 and 3573-5358. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1 \div 4^{25}$) times the increased probability for mismatch at each nucleotide position (3 x 25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant"(or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g., E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 90% sequence identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, and most preferably at least about 95% identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J. (1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The

term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1786 and 3573-5358 ; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1787-3572 and 5359-7144; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1787-3572 and 5359-7144. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1786 and 3573-5358 ; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO:1787-3572 and 5359-7144. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1786 and 3573-5358 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1786 and 3573-5358 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1786 and 3573-5358 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about 75%, at least about 80%, more typically at least about 90%, and even more typically at least about 95%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1786 and 3573-5358, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in

the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1786 and 3573-5358, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1786 and 3573-5358 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1786 and 3573-5358, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*,

hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1786 and 3573-5358, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1786 and 3573-5358 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1786 and 3573-5358 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following

vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for

transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1786 and 3573-5358, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID NO:1787-3572 and 5359-7144 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1786 and 3573-5358 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO:1-1786 and 3573-5358 , antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the

inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO:1-1786 and 3573-5358). For example, a derivative of a Tetrahymena L-19 IVS RNA can be

constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA

portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express

the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK,

HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5′ flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the

protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1787-3572 and 5359-7144 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1786 and 3573-5358 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1786 and 3573-5358 or (b) polynucleotides encoding any one of the amino acid sequences

set forth as SEQ ID NO:1787-3572 and 5359-7144 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1787-3572 and 5359-7144 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, typically at least about 95%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1787-3572 and 5359-7144.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, *e.g.*, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: *A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1787-3572 and 5359-7144.

The protein of the invention may also be expressed as a product of transgenic animals, *e.g.*, as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo, in situ,* or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular

endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus,

rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (*e.g.*, anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic

composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population.

Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention

may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide,

Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen

recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1- 7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such

transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol,* 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol,* 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem,* 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding

molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.*, increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications i.e. phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or

disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or

differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc.*, depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or

elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified

nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of

administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01μg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1μg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic

factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

5

4.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous,
10 intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

15 Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery
20 system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and
25 an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

30

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into
35 preparations which can be used pharmaceutically. These pharmaceutical compositions may be

manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers

enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with

an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well

known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable

lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions

may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which

modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the

population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the

invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as an amino acid sequence shown in SEQ ID NO: 1787, and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte

Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the

target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5.13.2 Monoclonal Antibodies

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.
The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego,

California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.
The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin

polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal

antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al,(Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 F_{ab} Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

5.13.5 Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the

binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to

stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular

defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon

a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (*e.g.* text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1786 and 3573-5358 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1786 and 3573-5358 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and

software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are

designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid

probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

4.18 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1786 and 3573-5358, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

4.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1786 and 3573-5358. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1786 and 3573-5358 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-$MeIm_7$), is then added to a final concentration of 10 mM 1-$MeIm_7$. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-$MeIm_7$, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *CviJI*, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *CviJI* normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (*CviJI***), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *CviJI*** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *CviJI*** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm^2, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm^2 and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5.0 EXAMPLES

5.1.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.1.2 EXAMPLE 2

Assemblage of Novel Nucleic Acids

The contigs or nucleic acids of the present invention, designated as SEQ ID NO: 3573-5358 were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

A polypeptide was predicted to be encoded by each of SEQ ID NO:3573-5358 as set forth below. The polypeptides was predicted using a software program called FASTY (available from http://fasta.bioch.virginia.edu) which selects a polypeptides based on a comparison of translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference. The predicted polypeptides are shown in Table 7.

5.2.2 EXAMPLE 3

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genebank. Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS:1- 327.

Table 1 shows the various tissue sources of SEQ ID NO: 1-327.

The nearest neighbor results for SEQ ID NO: 1-327 were obtained by a FASTA version 3 search against Genpept release 117, using FASTXY algorithm. FASTXY is an improved version of FASTA alignment which allows in-codon frame shifts. The nearest neighbor result showed the closest homologue for SEQ ID NO: 1-327 from Genpept . The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The nearest neighbor results for SEQ ID NO: 1-327 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the

signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 5 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.3.2 EXAMPLE 4

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 117, gb pri 117, UniGene version 117, Genpept release 117). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 328-1413.

Table 1 shows the various tissue sources of SEQ ID NO: 328-1413.

The nearest neighbor results for SEQ ID NO: 328-1413 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 118, using BLAST algorithm. The nearest neighbor result showed the closest homologue for SEQ ID NO: 328-1413 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in

the Sequence Listing. The nearest neighbor results for SEQ ID NO: 328-1413 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 5 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.3.2 EXAMPLE 5

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 117, gb pri 117, UniGene version 117, Genpept release 117). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide sequences, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 1414-1652.

Table 1 shows the various tissue sources of SEQ ID NO: 1414-1652.

The nearest neighbor results for SEQ ID NO: 1414-1652 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 118, using BLAST algorithm. The nearest neighbor result showed the closest homologue for SEQ ID NO: 1414-1652 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The nearest neighbor results for SEQ ID NO: 1414-1652 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 5 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.4.2 EXAMPLE 6

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 118, gb pri 118,

UniGene version 118, Genpept release 118). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide sequences, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 1653-1745.

Table 1 shows the various tissue sources of SEQ ID NO: 1653-1745.

The homology for SEQ ID NO: 1653-1745 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 118, using BLAST algorithm. The results showed homologues for SEQ ID NO: 1653-1745 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologues with identifiable functions for SEQ ID NO: 1653-1745 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 5 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.5.2 EXAMPLE 7

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 119, gb pri 119, UniGene version 119, Genpept release 119). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 1746-1768.

Table 1 shows the various tissue sources of SEQ ID NO: 1746-1768.

The homology for SEQ ID NO: 1746-1768 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 119, using BLAST algorithm. The results showed homologues for SEQ ID NO: 1746-1768 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologues with identifiable functions for SEQ ID NO: 1746-1768 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the PFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the PFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 5 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.6.2 EXAMPLE 8

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 120, gb pri 120, UniGene version 120, Genpept release 120). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.) . The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 1769-1786.

Table 1 shows the various tissue sources of SEQ ID NO: 1769-1786.

The homology for SEQ ID NO: 1769-1786 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and the amino acid version of Geneseq released on October 26, 2000, using BLAST algorithm. The results showed homologues for SEQ ID NO: 1769-1786 from Genpept. The homologues with identifiable functions for SEQ ID NO: 1769-1786 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by

reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 5 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

Table 6 is a correlation table of all of the sequences and the SEQ ID NOS.

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	9 19-21 50-51 65-66 72 78 80 82 85 87 107-108 113 116 123 138 140 150-152 159 169 177 192-193 202-203 212-214 225-226 235-236 251 258 268-269 272 280-281 295 298 301 321 326 331-332 334 356-357 362 369 379 382-383 416 423 443 459-460 473 475 477 488 496 500 503 519 526 547 574 582 587 608-609 613 618 633-634 645-646 652 657-658 660 669-671 678 687 695 697 710 715 724 731 775-777 796 804 811 857-859 862 869 899-900 912 919 922 924-929 933 936 962 979 988-989 996 1001 1004-1008 1018 1039 1047 1059 1064 1067 1070 1078 1082 1107 1113 1116-1117 1131 1134-1137 1140 1149 1151 1157 1180 1206 1229 1234 1241 1243 1258 1272-1273 1279 1288-1290 1294 1307-1308 1312 1320 1323 1330 1356 1360-1361 1368 1373-1375 1379 1391 1400 1417 1446 1468 1482 1493-1494 1501-1503 1506-1507 1512 1517 1522-1524 1530-1538 1537 1549 1565 1578 1598 1606 1608 1623 1625 1627 1639 1643 1648-1649 1653 1664 1667 1671 1696 1734 1741 1743-1744 1760-1761 1771
adult brain	GIBCO	ABD003	3 12-14 18-19 25 30-31 34-36 43-45 50-51 56 58 60 65-66 68-69 80 82 85 87 92 104 107-108 112-113 115-116 123-124 131-132 135-137 139 142 146 148-149 152 154 157 159 163 165 167 169 172 180 192-193 196-197 199 203 208 210 212-214 223 233 235-237 247 257 259 261 268-269 272 276 280-281 284-288 291-292 295 297 300-301 304 307 317 320-321 323 327 329-331 333-334 345-349 356-357 379-381 393 401 408 414 419 424 426-428 430 433-436 438-439 443 445 449 453-454 459-461 468 471-473 476-478 483 491 494 496 500 503 507-508 516 519-520 525-527 534 536-540 542-543 545 553 555 560 569-570 574-576 586-588 593 595 597 601 606-609 615-620 622-623 625 628-633 635-636 643 645-649 653 655-656 660-665 668-670 676 681 687 701 710 715 717 724-728 735 743 745-746 750 753 759 765-766 773 775-778 786 789 796 799-800 802-803 810-811 815 817 820-821 832 834-836 840 845-847 851 858-861 864 869 874 878 883 897 901-902 904-905 908 911-914 916 921-922 924-927 929 932-934 936-939 941-942 945 955-958 963 966-969 977 979-980 985-986 990 992-993 997-1001 1005-1007 1012 1017-1020 1023-1024 1029-1031 1034 1036 1039 1050 1059 1063-1066 1078 1081-1082 1085-1086 1089

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1097 1103 1107 1109 1112 1116-1117 1119 1121 1124 1127 1130 1134 1144-1145 1149 1151 1157-1158 1167 1170 1178 1184 1188 1190 1193-1194 1200 1202 1215-1217 1220 1226-1227 1229 1231 1241 1243 1247 1252 1258 1263 1267 1269 1279 1281 1284 1286-1289 1293-1294 1306-1307 1312 1316-1320 1326 1333 1338 1341 1344 1348 1351 1355-1357 1368 1374 1377 1380 1386 1389-1390 1394 1400 1409 1414 1422-1423 1425-1427 1437 1443 1446 1454 1456 1458-1459 1468 1470-1472 1478 1482-1483 1487-1488 1493 1497 1499 1506 1508-1511 1517 1522-1524 1530-1533 1545-1546 1548-1550 1552 1557-1559 1563 1565 1567 1569 1571 1586 1588 1591 1593 1595 1598-1601 1608 1611 1620-1621 1624-1626 1628 1630-1632 1636 1640-1641 1644-1645 1647 1649 1653-1655 1657 1664 1667 1669 1673 1678-1681 1686 1690 1694-1696 1701 1709 1711 1719 1722-1723 1726-1727 1731-1733 1738 1740 1743-1744 1747 1749 1753 1757-1758 1760-1761 1765 1771 1785
adult brain	Clontech	ABR001	9 29 68-69 113 115 146 152 206 223 245 277 307 320 324 330-331 344 348 352 362 379 384 393 404 408 414 441-442 454 469 481 490 506 517 586 597 631 641 659 691 715 799 803 833 865 871 875 880 882 908 920 937 1000 1005-1006 1027 1036 1041 1043 1075 1107 1112 1121 1127 1136-1137 1144-1147 1231 1238-1239 1280 1293 1320 1345 1355 1361 1383-1384 1400 1417 1448 1456 1476 1507 1570 1572 1609-1610 1614 1620 1626 1645 1653 1754 1759 1770 1786
adult brain	Clontech	ABR006	5-8 15-16 168 212-213 271 278 280-281 291-292 300-301 310 314 321 326 336-338 341 352 357 359-360 362 369 374 379 384 393 396-397 414 419-420 426-428 430 441-442 453 506 616-617 661 689 785 798 845 1018 1109 1113 1124 1148 1167 1187 1207 1227 1262 1265 1285 1312 1317-1319 1324-1327 1344 1369 1381 1400 1416 1421 1427 1430-1431 1436 1471 1501 1557-1559 1586 1588 1651 1653 1664-1665 1671 1673 1690 1697-1698 1700 1711 1717 1719-1720 1728 1736 1740 1743-1744 1757 1760-1761
adult brain	Clontech	ABR008	5-10 13-19 22-23 25 29 33 37-39 43-45 50-51 54-55 57-58 60-66 68-70 72 75 77-80 83 85 89-92 94 99-105 108-110 112-113 116-117 123 128 133 135-137 139 143 145-146 148 152 154-155 157 166 168-172 174-175 181-184 188-190 193-194 196 198-200 202 204-205 207-

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			208 210 214-215 218 221-226 229 231-232 234-241 245-247 251-253 255 257-259 268-269 271 276-281 285-286 288 290-292 300-302 304 307 309-311 313 315 317-318 320-322 325-326 328 330-331 333-338 341 344-347 349 352 354 356-357 362 369-373 376 379-380 382 384 387 390-391 393-394 397 399-403 405-411 414-415 417-420 426-428 437-438 440-444 453-455 462 464 467 469-471 476 478 482-484 488-491 497 503 506-513 516-517 520 524-526 528-530 532-534 537-540 542 544 547-551 553 561 565-567 572-574 577 581 585 587-588 590-591 597 599 601-602 606-610 612 615-617 619-620 622-623 628-629 631 633-634 636-641 643 645-647 651-653 655-664 669-671 673 679 682 687 689 691-700 702 706 710 715-717 720-721 725-734 736-739 742-743 746 750-752 756 758-759 762-764 766 768 773-778 780-782 784-785 787-789 794 796 799 802-803 805 811 814-815 818 825-826 834-837 839-840 842-843 856-859 861-862 865 867-872 874-875 881 883-884 887 889-892 894-895 897-898 901 904 908 910 912 914 917 919 921-924 926-927 930-932 935-941 943 945 949 953-954 958 961-963 967 969 971 975 977 981-983 986 988-990 992 997 999-1002 1004-1006 1008 1012 1018-1023 1027 1029-1031 1035-1037 1047-1048 1053 1057 1059 1063 1068 1070 1072-1075 1077 1081-1083 1085-1093 1095-1096 1108-1112 1114-1125 1127 1131-1133 1135-1138 1142-1145 1148-1158 1160-1163 1167 1169 1172 1175 1177 1180 1183-1188 1191-1195 1199-1200 1204 1206 1211 1213-1216 1222-1223 1226-1227 1229-1231 1234-1235 1241-1242 1244-1263 1266 1269-1271 1276-1277 1279-1281 1284-1286 1292 1294-1295 1299 1305-1309 1312 1314 1316-1319 1322 1324-1327 1330 1332 1334-1335 1339 1344-1346 1351 1354-1355 1357-1358 1365-1367 1369-1370 1373-1374 1376-1379 1381-1384 1386-1388 1392 1394 1396-1397 1400 1403-1407 1410 1414 1419-1420 1423 1432-1433 1435 1437-1438 1440-1442 1446 1448 1453-1455 1457 1461 1463-1464 1466 1468 1471 1477 1480 1482-1483 1496 1502-1504 1507-1509 1513 1519-1520 1524-1526 1536 1547 1549-1552 1567 1573-1574 1578 1586-1589 1597-1598 1601-1602 1605 1607-1609 1611-1617 1619-1621 1623 1625-1626 1635-1641 1643-1645 1649 1651 1653 1656-1658 1664 1669 1671-1674 1676-1684 1686 1689-1690 1694-1696 1704-1705 1708-1709

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1720-1724 1726-1728 1730-1733 1737-1740 1742-1745 1753 1756-1757 1759-1761 1765 1767 1771-1772 1776-1777 1779-1780 1786
adult brain	Clontech	ABR011	24 75 103 186 210 310-312 364-365 508 623 710 937 1002-1003 1059 1204 1609 1731-1732
adult brain	BioChain	ABR012	46 182-184 204-205 300 739 767 1371 1549 1620 1684
adult brain	Invitrogen	ABR013	185 204-205 364-365 393 497 595 687 692-694 830 845 1068 1320 1413 1640
adult brain	Invitrogen	ABR014	187 301 357 364-365 375 454 463 731 859 939 983 1073 1262 1270 1320 1403 1640 1651 1657 1696 1722 1738
adult brain	Invitrogen	ABR015	419 434-435 441-442 763 789 983 1320
adult brain	Invitrogen	ABR016	312 364-365 379 1320 1334-1335 1674 1722 1785
adult brain	Invitrogen	ABT004	14-16 22-23 25 37-39 43 58 60 70-72 78 86 94 107 113 116 136-137 143 146 152 161 173 182-184 194 196 198 210 218 229 259 267 295 298 309-310 320-321 324 336-338 346-347 349-350 356-357 362 371 379-380 382-383 391 393 396 399 401 408 428 438 459 461 476 482 490 502 507-509 516 526 531 557 562 597 602 607-609 624 652 655 667 669 671-672 687-689 695-696 710 712 715 721 732 739 743 750 753 766 778 780-781 789 803 814 826 830 837 841 857 869 874 894-895 925 937 949 954-956 960-961 963 968-969 988-989 1000 1005-1006 1016-1019 1021 1036-1037 1052 1086 1090 1109 1113 1115 1120-1121 1123-1124 1136-1137 1140 1144-1147 1151 1167 1170 1174 1188 1193-1194 1205 1225 1229 1231 1254 1258 1262 1280 1285 1309 1312 1334-1335 1341 1343-1344 1356-1357 1370 1378-1379 1383-1384 1403-1404 1423 1429 1434 1442 1448 1451-1452 1454 1470-1472 1482 1499 1525 1528-1529 1532 1536 1547 1554 1557-1559 1561-1562 1567 1585 1588 1590 1595 1601-1604 1608 1610-1613 1615 1619 1624 1627 1640 1644 1647 1660 1664 1666 1670 1675 1696 1704 1715 1723 1727 1738 1760-1761 1768 1779 1785-1786
cultured preadipocytes	Strategene	ADP001	5-8 11 17 25 68-69 80 82 87 103 105 110 116 136-138 168 171 188-189 196-198 261 267 276 288 293 301 318 331 336-338 379-380 391 400 428 430-431 510-512 520 524 527 549 557 561 602 618 620 622 631 637 647 670 681-682 710 731 748 782 793-794 817 834-836 843 845 858-859 879 882 893-895 934 960 982 986 995-996 1000 1002 1005-1007 1025 1027-1028 1032 1039 1045 1071 1078 1097 1099-1102 1136-1137 1140 1219-1220

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1260 1271 1297-1298 1314 1320 1322 1329 1339 1345 1365-1366 1370-1371 1398 1408 1423 1431 1437 1466 1468 1533 1539 1594 1602 1608 1614 1631 1649-1650 1660 1662 1673 1687-1688 1696 1711 1719-1720 1742 1746 1749 1760-1761 1765 1767 1771 1785
adrenal gland	Clontech	ADR002	4-10 15-16 25 29-31 43-45 47 50-51 55 60 62-63 65-66 75 80 102 116 118 122 126 130 137 150 169-170 181 192 198 201-203 215 227-228 247 251 255 267-269 271 280-281 285 295 298 311 336-338 342 349 351-352 354 372-373 383-385 391 400 410 415-416 424 426-427 431 434-437 439 445 454 461 473 477 483 491 493 497-498 503 516 519 527 535 546 549 552 572-573 581 588 595 600 602 608-610 620 628-630 637 645-646 670 679 703 713 715 719 732 734 744-746 758 773-778 789 816 829 837 845 848 869 875 883 898 904 912 922-923 930-931 942 948 952 965 967 969 976-977 981 990 992-993 1001 1004 1049 1055 1059 1071-1072 1076 1112-1113 1115 1121 1127 1134-1135 1151 1158 1163 1175 1181 1188 1209 1218 1224-1225 1227 1231 1243 1270-1271 1274 1280 1285 1290 1293 1307 1324-1325 1327 1330 1342-1343 1345 1348 1365-1366 1369 1378-1379 1387 1398 1400 1405 1417 1425-1426 1436 1440-1441 1444 1454 1463-1464 1488 1491 1507 1512 1538 1546 1567 1573-1575 1588 1598 1609 1614 1618 1622 1624 1627 1634 1636 1649 1651 1658 1671 1674 1678-1679 1691-1692 1703 1717 1727 1731-1732 1737 1765
adult heart	GIBCO	AHR001	4-8 10-11 15-16 18-21 34-39 44-46 50-52 57-58 60 62-63 71 75 82 85 87 89 94 97 100 103-104 108-110 112 114 116 118-119 122-123 127 130-132 134 136-138 141-144 147-151 153 163-164 168-171 179 186 192 195 197 199 204-205 212-215 220 225-226 229-230 232 234-236 251 257-260 262 265 272 274 277 280-282 285-286 289-292 296 298-301 304 307 309 314 321 324-325 330 333 336-338 345 349 351-352 354 358 361 368 370 380 383-384 387-388 391 393 397 401 406 408-409 411-412 414-416 430-431 433-439 445-446 449 452 454-455 457 459 462 469 472-473 476-480 483-484 487-490 492-493 496-498 503 506 508 510-513 516 519-522 526 534 536-540 542 546 549 553 560-562 574-577 581-582 584 586-587 589 593 595 597 604-609 611-612 615-620 622-623 626 632 637 645-652 656-660 665-666 670-672 674-675 683-684 687 692-694 697 701 709 712 715-716 719-720 725-

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			726 728 730-732 735 738-739 743-744 746 751 753 759 761 765 770-771 775-780 785 788-790 796 802 804 810 812 817 821 826 828 830 837 843 845-847 849-853 857-861 863-864 869 871 875 877-879 881 883 887 890-892 894-895 897-898 901 903 906-907 911-913 915 919 921-925 927-928 933-935 945 958 961-963 967 969-972 975 977-978 980-986 990 992 999-1002 1005-1007 1010 1016 1019-1020 1022-1023 1025 1028-1037 1039-1040 1043 1047 1050 1054-1055 1057 1059 1063-1064 1067-1068 1070 1072 1075-1076 1083 1085-1087 1089 1093-1094 1104 1106 1108-1109 1113 1116-1117 1119 1121 1124 1126 1128 1131-1134 1144-1145 1148-1149 1151 1158 1167 1169-1170 1175 1177 1192 1196 1199-1200 1202 1206-1208 1211 1216 1218 1222 1227-1229 1232-1235 1238-1241 1243-1244 1247-1248 1250 1253-1254 1256-1258 1261 1268 1270-1271 1277 1280-1282 1287 1292 1298-1299 1306 1308 1317-1321 1324-1325 1330 1332 1334-1337 1339 1344-1345 1349-1350 1354-1356 1359-1360 1365-1366 1369 1371 1374-1375 1378-1380 1383-1384 1389 1397 1400 1403 1409 1417 1423-1426 1437 1439 1442 1444 1446-1447 1450 1453 1468 1470 1473 1479 1481 1488 1490 1501-1504 1519 1521 1524 1528 1530-1534 1536-1537 1539, 1541-1542 1547 1553 1555 1560 1565 1567-1571 1588 1591 1597-1598 1601-1602 1605 1614-1616 1619-1620 1623-1628 1630-1632 1634 1636 1641 1644-1645 1647 1649 1652-1655 1659 1662 1667 1673-1674 1680-1681 1684 1686-1688 1704-1705 1709 1711-1712 1717 1724 1726-1727 1731-1733 1737-1738 1741 1743-1744 1749 1754-1755 1760-1761 1765 1772 1785
adult kidney	GIBCO	AKD001	4-8 10-11 17-21 29-31 35-39 42-45 50-51 56-58 60-61 64 68-69 75 77 80 82 85 87 92-94 97 100 102-104 107-108 112 116-117 119 123 127-133 136-137 139-141 143-144 147-154 157 161-163 165-166 169 172 176 178-179 192 194-197 199 201 203-206 209-210 212-213 215-216 223-228 234-236 238 247 251-253 257-259 261-262 265-269 271-272 274 276-277 279-281 284-286 290 293 296 298-299 301-302 304 307 311-313 321 325-326 329-331 333 341 344 348-350 352 356 358-359 362 364-365 368 370-372 374 376-377 380-382 392 395 398 400-401 404 407-409 414-415 423-424 430-437 443-444 446 449 451 453-455 459 461-462 464 467 469 471-474 476-477 480-481 483 487-488

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			490-491 493 497-505 510-513 516-520 522 524 526-529 534 537-540 544 547 549 554-556 560 562 564 567 571-576 578 582 586-589 592-593 598-599 601 604-606 608-613 615-619 621-626 632-634 637-643 645-652 655 660-664 669-672 676 678-679 688 692-695 698 702 711 713 717 719-720 727 731 735-736 738 743 745-746 751 753 755 762-763 765 771-773 775-778 780 786 788 793 795-796 800 803 805 808 810-812 814-819 821 826 829 832 834-838 842-845 848-855 857-861 864-865 867 869 871 874 876-883 886-887 889-891 893-896 898-900 902 906-908 910-914 918 920 922 925-927 929-935 937 940-942 945 948-949 951 953-958 960-961 963-964 969-970 972 976-978 982-986 988-990 992-993 995-997 999-1002 1004-1008 1010 1012-1013 1016-1017 1019-1020 1022 1025-1031 1035 1038-1040 1042 1044 1047 1050 1054-1055 1057-1064 1068 1070-1073 1078 1085-1086 1088-1089 1092 1094 1097 1099-1102 1107 1109-1112 1116-1119 1121 1123-1125 1132-1135 1140 1142-1143 1146-1147 1149-1150 1153-1154 1157 1159 1163 1167 1170 1178-1179 1181 1183 1192 1196-1200 1202-1204 1206-1211 1216-1219 1221-1222 1225 1227-1230 1232-1234 1238-1241 1243-1244 1246-1247 1253 1257-1258 1260-1261 1267-1268 1270 1272-1274 1281 1283 1287-1289 1293-1295 1299 1306 1308 1311-1313 1317-1320 1323 1329-1330 1334-1335 1339 1341 1349-1350 1353-1357 1359 1367 1369 1373 1375 1378-1379 1394 1397 1400 1403 1405 1407-1409 1417 1419 1423-1424 1428-1431 1433 1437-1438 1442-1443 1445-1446 1448-1450 1453-1454 1459 1461 1465-1468 1474-1475 1478 1484-1488 1490 1492-1493 1495 1497-1498 1506-1507 1509 1512 1518 1521-1522 1525 1527-1528 1532-1533 1537 1540-1541 1547-1550 1552 1556-1559 1561 1565-1566 1568 1571 1575 1578-1579 1583 1586-1587 1589 1591-1592 1594 1598 1600 1603-1604 1606 1608 1611 1613 1615-1616 1618-1622 1624-1628 1631-1632 1634-1636 1638-1639 1641 1644 1646-1649 1653-1656 1662 1664 1666-1667 1670-1671 1676-1679 1683-1684 1686 1691-1692 1696-1699 1701 1709-1711 1713-1714 1716-1719 1723-1724 1726-1727 1733 1737-1738 1741 1743-1744 1748-1749 1751 1760-1761 1763-1768 1778 1780 1785
adult kidney	Invitrogen	AKT002	20-21 37-39 47 52 57 60 65-66 68-69 80 104 107-108 122 130 133 136-137 140 142-143 149 169 174

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			181 197 227-228 235-236 244 251 261-265 267 280-281 286 290 299 301 304-305 309 312-313 339 341 344-345 349 358 370-372 376 382-383 387 392 401 414 416 421 430 443 445 449 453-454 472 487-488 504 506 513 516 519 522 528 536-540 546 554 585 587 594 598 602 607 616-617 626-627 636 643 662-664 695 709 721 735 743 761 768 775-777 788 796 804 814 827 837-838 849-850 852-853 869-870 881 890-892 898 903 905-907 914 919 925 927 934 941 949 952 957 960 962 968 970 1000 1008 1029-1030 1044 1052 1055 1063 1067-1068 1073 1085 1099-1102 1107 1110-1111 1113 1115 1119 1126 1134 1136-1137 1146-1148 1153 1159 1192 1196 1199 1232-1233 1241 1256 1264 1272-1273 1281 1285 1293-1294 1299 1312 1320 1324-1325 1330 1344 1349 1351 1355-1356 1369 1378-1379 1403 1414 1419 1428-1429 1436 1446 1458 1463-1464 1467-1468 1470 1477-1478 1486 1491 1509 1519 1527 1529 1534 1547 1596 1600 1619 1623 1629 1631 1634 1638 1643 1647 1652 1660 1664 1667 1669-1670 1673 1686 1709 1727 1740 1776
adult lung	GIBCO	ALG001	4-8 14 37-39 44-46 50-51 56 62-63 75 82 88 93 103-104 113 125 133 140 143 150 152 154 157 162 171-172 174-175 190-191 196 200 211 214 219 223-224 227-228 251-252 256 265 272 274 280-281 285 310 332 345 351 362 371 381-382 394 408-409 431 436 445 454 459 461 467 469 471 476-477 488 504 513 527 537-540 544 547-548 554 564 583 607 616-617 621 623-624 634 645-646 662-664 670 695 716 719 743-744 763 766 774 789 803 811 814 817 831-832 837-838 845 852-853 858-859 861 866 880 887 901 905 941 954-957 966 971 977 979 981 987 990 992 996 1001 1005-1006 1014 1017 1045 1047 1054 1059 1062 1064 1072 1080 1086-1089 1094 1107 1126 1134 1136-1137 1142 1150 1157 1173 1190 1200 1208 1220 1241 1272-1273 1280 1282 1295 1306 1320 1331-1332 1353 1374 1379 1383-1384 1404 1409 1423 1434 1436 1442 1474 1478 1494 1509 1522 1525 1531-1532 1547 1549 1553-1554 1571 1598 1606 1613 1624 1627-1629 1632 1642 1644 1662 1669 1676-1677 1684 1696 1727 1731-1732 1737-1738 1748-1749 1786
lymph node	Clontech	ALN001	4 24 50-51 82 105 137 153 198 201 223-224 234 268-269 272 280-281 287 301 312 329 343 382 421 430 433 445 451 461-462 475 481-482 503 526 529 537-540 546-547

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			621 626 649 679 719 725-726 738 793 803 831 834-836 838 844 857-858 866 879 905 913 928 963 976 1005-1006 1012 1038 1050 1116-1117 1151 1199 1204 1226 1243 1265 1274 1324-1325 1339 1353 1374 1377 1440-1441 1447 1504 1549 1600 1618-1619 1631 1641 1644 1653 1687-1688 1691-1692 1741 1771
young liver	GIBCO	ALV001	5-8 11 20-21 46 50-51 58 65-66 75 79 82 93 97 102-103 108 110 116 139 143-144 148-149 171-172 174 187-189 194-195 198 209 214-215 230 250 258 267-269 280-281 306 309 342 351 356 359 362 372 374 392 394 398 401 407-408 410 414 431 444 455 459 476 478 483 493 510-512 516 520 522 526 536 549 571 574-577 585 592 601-602 607 621-624 628-630 632-633 637 648 660 666-667 678 697-698 700 717 719 728 730 734 738 744-745 766 770 773 779 788 800 808 812 814 841 849-851 871 874 879 887 893 898-900 902-904 906-907 911 919 922 924 934 953 957 963 965 970 984 986 997 1001 1004 1007 1012 1029-1030 1033-1034 1052 1061 1066 1070 1076 1086 1089 1093 1099-1102 1110-1112 1116-1117 1119 1121 1125 1136-1137 1144-1145 1156-1157 1159 1196 1199-1200 1209 1211 1219-1220 1241 1244 1262 1270 1275 1279 1283 1295 1317-1320 1332 1339 1344 1359 1362-1363 1379 1383-1384 1403 1415 1430-1431 1437 1450 1467 1475-1476 1483-1484 1494-1495 1498 1505 1512 1516 1518-1519 1526 1529 1547 1550-1552 1557-1559 1565 1583 1587 1597 1609 1614 1620 1631 1637 1641 1644 1654-1655 1662 1667 1669 1684 1691-1692 1702 1711 1725 1738 1741 1743-1744 1758 1760-1761 1763-1765 1769
adult liver	Invitrogen	ALV002	5-8 17 20-21 32-33 41 55 58 64 75 77 86 89 102 108 117 119 175-176 198 200 209 231 235-236 250 272 275-276 284 306 316 321 325 333 356 359 374 376 398 401 408 414 428 430 433-435 454 476 494 503-505 517-518 528 534 544 552 561-563 567 578 581 608-609 630 632 637 644 650 661 665 672 702 707 710 721-722 750 753 778 782 794 814 820 826 834-837 847 849-850 858 861 874 879 893 898 904 911 918 921-922 926 946 948 972 978 986 996 1020 1027 1031 1034 1053 1063 1068 1070 1073 1086 1089 1093 1097 1113 1119 1156 1159 1195 1198-1199 1208 1220 1227 1241 1261 1272-1273 1277 1285 1308 1315 1320 1324-1325 1330 1362-1363 1375 1403 1408-1409 1415 1431-1432 1435 1467 1469 1482 1504 1524 1542 1547

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1550 1567 1578 1581 1583 1594 1597 1601-1602 1611-1612 1615 1618-1619 1621 1625 1637 1645 1647 1652 1654-1655 1660 1666 1669-1671 1684 1706 1722 1737-1738 1742-1744 1760-1761 1753-1765 1772 1774
adult liver	Clontech	ALV003	29 676 997 1063 1119 1536 1766
adult ovary	Invitrogen	AOV001	1 4-18 20-23 29 35-40 42-48 50-51 53-58 61-63 65-66 68-69 73-75 77-78 80 82 85 87 89 97 100-101 103-104 106-108 110 113 115 118 122-124 126 128 133-134 136-140 142 145-147 149-157 161 166 168-170 174 177-178 180 182-186 188-189 192-203 207 209 211-215 219 221-224 229-230 234 242-243 246-247 255 258 260-262 265-269 271-272 274 277-281 284-286 288 290 295 299 301-302 304 307 309-311 313-314 316 321 323-326 330 332-333 335-338 341 344 349 352-353 356 358 360 362 370-372 376-377 379-384 387 390-392 394 397-398 400 403 408-410 412 414-416 423-424 426-427 430-435 439 443-446 448-449 451 453-455 462-463 468-471 473 476-479 481-484 487 489-494 496-497 499-501 503-505 509-514 516-517 519-520 522 524 526 528-534 541-544 546-547 549 552 554-555 561-564 566-567 569-570 572-573 575-576 579 581 583 585-588 590-591 593 595 597 599 601-605 607-613 615 618-622 624-627 630 632-633 636-640 642 644-647 649-652 654-655 657-665 667-675 677-678 681 683-684 692-695 697-710 714-721 723 725-727 729 732 734-735 743-746 750-751 753 758 763 765 767 772-773 775-778 780 783-784 786 788 790-791 794-796 800 803 805 809-811 813-815 818-819 821-824 826 828-829 831-832 837-838 843-850 852-857 859-864 867 869 871-872 874-875 878-883 887-888 890-895 898-910 912-914 916 919-922 924 926-927 929-939 941 943-946 948-951 953 955-958 961-964 966-967 970-979 981-982 985-986 988-990 992 995-997 999-1001 1004-1009 1011-1013 1016 1019-1020 1024-1025 1029-1031 1033-1035 1037 1039 1041-1047 1050-1051 1054-1060 1062-1064 1067-1070 1072-1073 1075-1076 1078-1079 1085-1086 1089-1090 1094-1096 1098-1103 1106-1108 1112-1117 1119-1120 1123-1127 1131-1135 1142-1143 1146-1149 1153 1156 1158 1163 1165-1166 1169-1171 1173-1175 1177-1178 1180 1183-1185 1190-1191 1195 1197-1200 1202 1205-1214 1217-1219 1221-1226 1232-1235 1238-1241 1243-1244 1247 1249 1252-1254 1256-1258 1262 1265 1267-1268 1270 1275 1278 1280-1283 1286-1289 1291 1293-1294 1298-

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1299 1306 1308 1312 1317-1321 1323 1327 1329-1330 1332-1333 1338-1339 1341 1343-1351 1356 1359 1361 1365-1366 1371-1375 1377-1379 1383-1384 1386 1389 1394 1400 1404 1416-1417 1422-1427 1429-1431 1435-1436 1439-1443 1445-1450 1453-1454 1459 1463-1464 1466 1468 1470 1474-1481 1484-1485 1488 1491 1493-1494 1496-1498 1501-1504 1506-1507 1511-1517 1519 1521-1524 1526-1527 1530-1531 1534-1536 1538-1539 1541 1546 1548-1550 1553 1555-1559 1561-1563 1566-1567 1569-1570 1572 1574-1575 1578 1580-1581 1587-1588 1590-1591 1595 1597-1598 1600-1606 1609 1611-1621 1623-1630 1634 1636 1638 1641 1643 1645 1647-1657 1659-1662 1664 1667 1669-1671 1673-1674 1676-1681 1683-1690 1699 1702-1707 1710-1711 1713-1714 1716-1719 1723-1724 1726-1728 1731-1733 1735 1737-1738 1740-1741 1743-1744 1748-1751 1753 1755-1756 1760-1762 1765 1767-1768 1770-1771 1776 1778-1779 1783-1784 1786
adult placenta	Clontech	APL001	5-8 44-45 90-91 107-108 159 178 311 351 414 476 503 545 574 624 636 719 755 773 860 890-891 924 947 955-956 962 990 992 1002 1045 1202 1320 1369 1628 1686 1713-1714 1743-1744
placenta	Invitrogen	APL002	14-16 26 29 43 60-61 79-80 103 106 116 135 171 177 180 194 196 198 210 216 235-236 272 290 299 309 329 334 339 359 379-380 417 423 430 434-435 448 454 483 490-491 517 522 631 723 725-726 728 738 746 769 818 843 854-855 857-858 916 948 953-954 976 988-989 1005-1006 1013 1033 1036 1064 1068 1070 1086 1139 1144-1145 1160 1277 1285 1317-1320 1343 1345 1429 1435 1438 1454 1482 1486 1490 1512 1519 1532 1549 1592-1593 1602 1626 1647 1649 1664 1673 1675 1722 1727 1730 1746 1776
adult spleen	GIBCO	ASP001	3 5-8 12 15-16 19-21 24 29 34-36 44-45 57 60 82-83 87 89 94 98-99 103 106 108 117 119-121 139 141 147 152-153 155 166 169 171 174 178-180 196 198 201-206 209-211 215 219 234 253-254 256 258 264 272 280-281 290 295 302 309 312 325 333 341 349 358 372 382 386-387 394 406 414 431 434-436 446 448 451 473 481 490-493 500 503 505 517 519 530 534 536-540 547 554 557 574-576 582 592 595 604 611-612 620-621 623 631-632 642 652 659 661 667 671 673-675 684 700 721 728 730 732 738 742-744 746 762 765 774 780 788-789 794 810-811 817 822 830 832 845 848 852-853 858 862 866 874 879 882

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			884 906-908 912 919 921-923 926-927 934 942 949 957-958 963 977-978 983 990 992-994 996-997 999 1005-1007 1010 1012 1031 1036 1042-1044 1046 1049 1059 1068 1070 1076 1089-1090 1094 1103 1109 1113 1115 1124 1140 1163 1170 1174 1177 1190 1196 1219-1220 1226-1227 1229 1236 1241 1246 1258 1269 1271 1274 1295 1301 1320 1322 1330 1334-1335 1339 1349 1351 1353 1359-1360 1364 1369 1374 1386 1397 1413 1417 1434 1436-1437 1439 1468 1474 1477 1480 1485-1487 1498 1512 1522 1525 1544-1549 1553 1560 1567 1591 1600 1631 1636 1651 1654-1655 1658 1662 1670 1674 1678-1679 1684 1686 1700 1727 1733 1738 1740-1741 1760-1761 1774 1779 1781-1782
testis	GIBCO	ATS001	5-8 10 26 30-31 47 50-51 57 68-69 82 84-85 97 102 113 119 137 139 150 152 154 156 163 169 174 176-177 192 194 196-197 212-215 227-228 247 255 258 261 282 285 288-289 301 307 311 316 330 334 349 370-372 392 398 410 415 426-427 430-431 433 437 446 454 461 469 473 477 481-482 493 499 502-503 513 522 526 547 552-553 563-564 572-573 575-576 581-582 585 599-602 605 612 615-617 620 631 637 647 649-650 656 660 665 670 674-675 712 719-721 723 728 731 738 744 746 773 780 784 788-789 802 804 809 811 814 826 831 837 843 845 848 859 866 869 877 905 913 916 919 921 926 929 937 950 960 963 971 975 977 981 990 992-993 1007 1016 1029-1030 1034-1035 1038-1039 1045 1059-1060 1064 1070 1072-1073 1087 1089 1097 1099-1102 1104 1108 1113 1141 1149 1161-1162 1175 1208-1209 1222 1227 1229 1231 1235 1238-1239 1243 1253 1285 1287-1289 1291-1293 1307 1311 1317-1320 1330 1332 1338 1345 1369 1373-1374 1379 1389 1399-1400 1409 1423-1424 1430 1435-1437 1443 1459 1484 1486 1490 1493 1496-1497 1501 1505 1509-1513 1527 1530-1531 1533 1537 1546 1549 1563 1565 1567 1569 1571 1577 1586 1591 1599 1602 1625 1628 1630-1632 1636 1639 1642 1649 1661-1662 1666-1667 1670 1675 1684 1690 1699 1705 1712 1717 1724 1730 1737-1738 1752 1767 1779
Genomic DNA from BAC 63I18	Research Genetics (CITB BAC Library)	BAC001	686 1352 1412
Genomic DNA from BAC 393I6	Research Genetics (CITB BAC Library)	BAC002	1411-1412

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
Genomic DNA from BAC 393I6	Research Genetics (CITB BAC Library)	BAC003	1352
adult bladder	Invitrogen	BLD001	5-8 17-18 22-23 33 37-39 56-57 80 93 100 120-121 169 201 237 251-252 272 278 311 348 363 382 413 415 424 430 443 483 502 542-543 562 564 607 616-617 626 635 652 667 671 710 727 755-756 762 773 786 789 837 840 866 893 898 909 918 929 966 977 983 1016 1025 1055 1073 1082 1140 1167 1185 1189 1199 1270 1369 1481 1536 1560 1573 1596 1614 1636-1637 1649-1650 1654-1655 1658 1669 1671 1690 1719 1727 1731-1732 1739 1741 1760-1761 1779
bone marrow	Clontech	BMD001	3-8 11 13 18 29-31 33 35-36 40 43-45 47-48 50-51 57 60 65-66 75 80 82 85 88-89 94 100 103 107 110 115 118-119 124-125 133-134 136-137 139-141 146 150 152-153 155 161 163 168-170 172 178-180 187 192-193 197-198 203-205 210-213 215 217 219 222 224-226 233 235-237 242-244 255 258 260 263-264 266 273 276 278 283 286 290 295 301-302 307 312-313 321 330 333 339 343 352 357-358 370-371 382 384-385 387 389 394 408 410 412 416 421 424-427 429-431 436-437 439 441-442 445 447 454-456 461-462 471-472 475 477-479 481-482 485 488 493 498 500 503-506 513 516 519 523-524 526 530 535-540 542 544-545 549 555 565 567 569-577 581 583-586 588 593 601 603-604 608-609 613-619 621-622 632-633 636-637 642 649-650 656-660 666 670 672 674-675 679 683 701 708 716 718-720 731 735-736 740-742 744-745 752 761 765 772-773 775-778 780 785-786 789-791 796 798 802 810-812 823-824 826 830 832-833 837-838 843-844 848-855 858-859 866-867 869 878-880 883 890-892 896 903 905 908 912-914 922-924 927 930-931 937 939-941 952-953 955-958 963 969 973 976 981 985 987 990 992 995 1000 1002 1005-1007 1013 1016 1025 1028-1031 1033 1035 1037 1039 1042 1044 1047 1050 1053-1054 1059 1061 1063 1066 1070-1071 1079 1106 1110-1113 1115-1117 1124 1126 1134-1135 1142 1144-1145 1163 1172 1178 1197 1199-1200 1202 1216-1217 1224 1227-1228 1240 1246 1254 1261 1266 1270 1278 1281 1285 1287 1290-1291 1293 1299-1301 1308 1314 1317-1320 1327 1331 1339 1343 1346 1349 1353 1356 1361 1367 1369 1372-1374 1379-1380 1394 1400 1403 1406 1408 1413 1417 1419 1423 1425-1427 1430-1431 1433 1439 1443 1446-1449 1459 1463-1464 1482 1486 1493-1494

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1506 1509 1513 1521-1522 1524 1526 1528 1531 1536-1537 1543 1546 1548-1549 1552 1554-1555 1557-1559 1571-1572 1581 1589-1592 1597-1600 1609 1614 1621 1626-1628 1630-1632 1634 1636 1638-1639 1641 1646-1647 1651 1653-1655 1661-1662 1676-1681 1684 1686 1690 1702 1707 1711 1713-1714 1717 1720 1722-1723 1727 1737-1738 1740 1758 1767 1772 1781-1782 1785-1786
bone marrow	Clontech	BMD002	11 15-16 19 30-31 35-36 68-69 75 83-84 93 99 103 108-109 118 137 139 169-170 174 177 180 190 193 212-213 219 222 225-226 232 237 255 259 264 273-274 284 286 290-292 295 301 303-304 307 312-313 316 324 326 330 334-335 348 352-353 357 360 370-373 384 386-387 397 403-404 414-416 421 425-427 429-430 433-436 440 444 451 454 465-466 472 475 478 491 493 516 520 523 525 531 545 548 552 566 569-570 581 583 590-591 597-598 601 616-617 621 641 650 652 656 659 671 674-675 679 684 710 718-719 728 734 737-738 742 761 765 774-778 790 811 814 818 830 834-836 854-855 859 866 869 871 878-879 884 889 892 904 922-923 932 990 992 998 1001 1004 1016 1036 1042 1048 1051 1054-1055 1058 1088-1089 1106 1112-1114 1155 1157 1192 1200 1223 1227-1228 1236-1237 1260-1261 1282-1283 1285 1287 1295 1314 1317-1321 1324-1327 1330 1333 1341 1343 1347 1350 1353 1355-1357 1367 1369-1370 1373 1377 1379 1381 1383-1384 1394 1397 1400 1406 1413 1417 1425-1427 1438 1442 1446 1459-1460 1470 1493 1505 1521 1536 1546-1549 1560 1573-1574 1578 1598-1600 1621 1626 1631 1634 1646 1649 1653 1656 1658 1669-1670 1683-1684 1687-1688 1690-1693 1696 1699 1702 1704 1707-1709 1711 1720 1722-1723 1725 1727 1729 1731-1733 1738-1740 1743-1746 1752 1755 1760-1761 1767 1777 1781-1782 1786
bone marrow	Clontech	BMD004	73-74 503 922 1036 1711
bone marrow	Clontech	BMD007	95-96 866 1320 1475
adult colon	Invitrogen	CLN001	17 56-58 103 110 117 144 150 171 179 185 188-189 201 204-206 210 218-221 225-226 231 237 251 277 288 310 312 320 333 359 386 388 394 408 420 455 481 485 503 510-512 590-591 615 635 647-648 665 672 684 697 710 725-726 743 780 786 788 826-827 848-850 854-855 858 866 872 898 918 921-923 953 976 983 993 1005-1006 1017 1020 1025 1027 1054-1055 1063 1068-1069 1140 1153 1170 1185 1196 1199 1220 1280 1314-1315 1320 1345 1351 1355 1369 1428 1439

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1462-1464 1512 1556 1583 1587 1594 1596 1614 1625-1626 1631 1639 1645 1650 1675-1677 1687-1688 1701 1713-1714 1724 1740 1765
Mixture of 16 tissues - mRNAs	Various Vendors	CTL016	401 1490 1686
Mixture of 16 tissues - mRNAs*	Various Vendors	CTL021	312 782 1132-1133 1403 1712 1715
adult cervix	BioChain	CVX001	1 4-8 11 13 18-21 25-26 30-31 33 37-39 43 46-47 58 61 64-66 71 73-74 82 85 94 100 103-104 113 118 122 126 130 134 140 147 153-156 163 170 179 181 186 192 195-196 198 201-202 218-219 222 229-231 257 266 276-277 285-286 288 298 301-302 304 307 312-314 324 326 329-330 332 335 342 352 358 362 371-372 376 379 381-382 384 388 398 400 410 414 416 419-420 426-427 430-431 433-436 439 446 448 461-462 464 471-477 479 482-483 491 493 496 503 506 510-513 516-517 526 530 535 542-544 546-547 557 561 572-573 575-577 581-582 585-586 588-589 593-594 600 602 604-605 607-609 612 615-619 623 644 650 654 657-658 662-665 670 672 680 683 691-694 698 706 708-709 711 713 720-721 727 729 731-732 737 745-747 753-754 760 765 771 774-777 780 790 793 796 798 800 803 805 818 826 828 831-832 834-836 843 847-848 851-855 857-860 864-866 869 871 876 878-880 882 887 890-891 897 899-902 905-908 912-913 916 918-919 922 927 932 934-938 944 948 955-956 958 963-964 967 969-970 972 976 978-979 983 985 990 992 1000 1005-1007 1016-1017 1024 1027 1033 1036 1038 1045 1047 1053-1056 1066-1067 1071 1073 1075 1079 1082 1098 1113 1124 1129 1134 1139 1146-1149 1163 1167 1170 1173 1175 1177 1181 1197 1200 1202 1211 1214 1216 1221-1222 1225 1227 1232-1234 1240-1241 1243 1258 1264-1265 1268 1270 1279 1287-1290 1308 1310-1311 1316 1320 1323 1327 1345 1349 1353-1354 1360 1372-1374 1383-1384 1386 1394 1397 1405-

* The 16 tissue-mRNAs and their vendor source, are as follows: 1) Normal adult brain mRNA (Invitrogen), 2) normal adult kidney mRNA (Invitrogen), 3) normal adult liver mRNA (Invitrogen), 4) normal fetal brain mRNA (Invitrogen), 5) normal fetal kidney mRNA (Invitrogen), 6) normal fetal liver mRNA (Invitrogen), 7) normal fetal skin mRNA (Invitrogen), 8) human adrenal gland mRNA (Clontech), 9) human bone marrow mRNA (Clontech), 10) human leukemia lymphablastic mRNA (Clontech), 11) human thymus mRNA (Clontech), 12) human lymph node mRNA (Clontech), 13) human spinal cord mRNA (Clontech), 14) human thyroid mRNA (Clontech), 15) human esophagus mRNA (BioChain), 16) human conceptional umbilical cord mRNA (BioChain).

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1406 1416 1425-1427 1431 1436-1437 1442 1446 1448 1453 1459 1466 1472 1478 1482 1496 1501-1503 1506 1512 1522 1527-1528 1531 1533 1541 1547 1569 1571 1585 1589 1597-1598 1600 1608-1609 1614-1616 1620 1623-1624 1626-1628 1630 1638 1641 1643 1649 1653 1656 1662 1667 1669 1674-1675 1683 1685-1688 1699 1702 1709-1710 1715 1717 1722 1724 1729 1731-1732 1735-1739 1741 1743-1744 1748-1749 1755 1760-1762 1767 1773 1778 1785-1786
diaphragm	BioChain	DIA002	137 282 289 730 780 986 1409 1478 1599 1614
endothelial cells	Strategene	EDT001	3 5-10 13 15-21 24-26 29 34 37-39 42 44-45 50-51 53-55 57-58 60-61 65-66 68-69 73-74 77-78 80 82-83 85 87 89 93-96 101-105 108 110 112-114 116 118-122 124 128 133-134 137-142 147-150 152-153 161-163 166-172 176-179 187 190 192 194 196-201 204-207 210 212-214 220 224 229-230 233 235-236 240-241 251-252 258 261-262 265 267-269 272 276-277 279-281 284-285 288 290 295-296 301-302 310-311 313 316 321 325 329 331-333 335 340-342 351-355 360 371 375 380-382 384 387 390 392 397 400 407-408 410 412 414 416 425-427 431 434-436 439 444-445 449 454 463-464 472-475 477-479 486 488-490 497-498 500-504 510-513 516-519 522 524 526-528 532-534 536-540 542-546 548 561-563 566-567 572-576 579 581 585-586 589 593 595 597 599 603 607-612 615-617 620 622 626 630 632-634 638-641 644 647 656-660 662-664 670 673 678 680-682 692-697 707 709-710 712-713 719 730 732 734 736 738 743-746 751 759 768 771 773 775-778 783 786-789 793 800 803 805-807 810-811 814 816-818 821-822 824 826 828-829 832 834-838 842-845 848-850 854-860 862 864 869 871 874 876-879 883 885 887 890-891 894-895 898-900 903 908 910-913 916 919-922 924 926-928 930-935 939 943 948-949 951-954 957 959-961 964 969-970 973 975-978 983-984 988-990 992-993 996-997 1000 1002 1004-1013 1016-1020 1022-1025 1028 1031 1033-1034 1038-1046 1050 1055-1056 1059-1060 1062-1064 1067-1070 1072-1074 1076 1078 1082 1086-1087 1089-1090 1093-1097 1099-1103 1107 1109-1113 1116-1117 1124-1126 1128-1131 1134-1135 1138 1140 1144-1145 1148-1149 1153 1157 1160 1163 1171 1183-1184 1198-1199 1202 1205-1207 1211 1216-1217 1219 1221 1225 1229 1232-1235 1238-1241 1243-1244 1246 1250 1253 1257-1258 1261

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1265-1266 1268 1270-1271 1274-1277 1280-1283 1285-1286 1288-1290 1293 1295 1298 1308 1312 1317-1320 1324-1325 1327 1329-1330 1334-1335 1338 1342-1343 1345-1347 1350 1355-1356 1359 1367 1369 1374 1376 1379 1398 1400 1406 1408 1414 1417 1419 1424-1426 1428-1431 1434-1438 1440-1442 1448 1450 1462-1466 1468 1472 1474 1478 1487-1488 1491-1493 1501-1504 1506 1509 1511 1516 1520-1521 1526 1529 1531 1536-1537 1539-1540 1546-1547 1549 1552 1555 1557-1559 1561-1565 1568 1571 1575 1578-1579 1581-1583 1587-1588 1590 1592 1597 1605-1606 1611 1613 1615 1618-1621 1624-1628 1630-1631 1634 1636 1638 1641 1643-1650 1652-1659 1664 1666-1667 1669 1671 1675-1681 1683-1688 1696-1698 1703 1711 1715-1716 1719 1722-1723 1726 1731-1733 1736 1739-1741 1743-1744 1749 1755 1760-1761 1765 1767-1768 1771-1773 1776 1779 1783-1786
Genomic clones from the short arm of chromosome 8	Genomic DNA from Genetic Research	EPM001	286 686 1297 1303-1304 1352 1411-1412 1754
esophagus	BioChain	ESO002	131-132 261 289 380 503 860 892 1000 1007 1397
fetal brain	Clontech	FBR001	62-63 89 112 126 194 322 336-338 379 391 411 481 546 563 607 679 710 867 1012 1031 1055 1251 1262 1320 1407 1643 1652 1686 1731-1732 1746 1765
fetal brain	Clontech	FBR004	68-69 90-91 139 212-213 301 331 362 374 403 436 611 645-646 659 668 670 691 785 805 845 1163 1209 1216 1232-1233 1238-1239 1387 1410 1416 1430 1496 1536 1547 1593
fetal brain	Clontech	FBR006	5-9 25 43 60 62-63 65-66 70 72 80 87 92 101 103 108 114 136 139 149 152-153 157 168 171-172 175 207-208 210 212-213 221-226 237-238 251-253 266 272 279-281 295 301-302 307 310 317-318 321-324 330 333-334 336-338 346-347 352 357 370 373 377 379-380 382 384 391-392 397 399 402 406-408 410-411 417 421 424 426-427 430 436-437 440-443 454 460 464 467 473 476 483 488-489 495 497 508 510-513 516 519-520 524 530 537-540 544 547 550 561 567 572-574 582 590-591 595 597 604 607-609 615 623 628-629 631 634 638-640 655 657-658 660 665 669 674-675 679 689 691-694 696-697 699 701 706 710 716 720 728 732 734 736 742-744 757-760 763 775-778 780 799 806-807 810 817-818 826 839 843 858 861 864 871-872 884 890-891 894-895 898 904 915 921-923 935-936 938 945 950 952 955-956 958-959 961 963 967 969-971 990 992

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			999 1001 1005-1006 1008 1013 1016 1022 1024 1029-1030 1032 1035 1042 1047-1048 1052 1056 1065 1067 1070 1082 1089 1109 1114-1115 1119 1131 1143-1149 1151 1153-1156 1160 1163 1167 1172-1173 1178 1184 1186 1188 1190-1200 1211 1216 1222-1223 1226-1227 1229 1231 1236 1245 1253-1255 1258 1260 1262 1266 1270-1273 1281 1287 1308-1309 1314 1317-1320 1326 1334-1335 1339 1341 1344 1350 1356 1369-1371 1373 1376 1379 1381-1382 1386 1392 1396-1398 1419 1423 1425-1426 1428-1429 1432 1437 1440-1441 1448 1466 1470 1482 1502-1503 1507 1511 1513 1526 1519 1536 1544 1549-1550 1557-1559 1573 1589-1590 1598 1608 1611-1614 1619 1621 1625-1626 1640 1651 1657-1658 1676-1679 1693 1696 1703-1704 1713-1714 1718 1720 1722 1724 1726 1728 1730-1733 1735-1736 1738-1739 1742 1745 1755 1759-1761 1765 1767 1771-1772 1777 1779-1780 1786
fetal brain	Clontech	FBRs03	235-236 520 864 1068 1188 1587
fetal brain	Invitrogen	FBT002	15-18 20-21 24-25 29 34 43 61-63 77-78 98 101 103 107-108 128 130 136 146 148 165-166 171 174 181 185 196-198 204-205 208 223 230 235-236 251 253 261 268-269 280-281 284-285 288 309-311 321 329 334 339 346-347 350 357-359 381-383 390 407 418-419 430 434-435 438 443-444 461 464-466 483 490 494 509 516 519 522 527 557 561-562 572-573 590-591 595 597 623 632 647-648 650 655 669-670 672 682 690-691 700-701 710 717 736 746 782 784 788-789 814-815 825 829 840-841 847 854-855 857-858 897-900 904 919 925 935-937 946 948-949 954 960-962 966 969-970 986 996 1000-1001 1005-1007 1012 1014 1022-1028 1045 1052 1055 1068 1070 1072 1078 1082 1085 1090 1109 1115 1118 1120 1128 1136-1137 1144-1145 1149 1156-1157 1193-1195 1198 1204-1205 1220 1222 1234 1257 1262 1271 1274-1275 1280 1285-1286 1294 1312 1314 1317-1320 1330 1342 1344-1345 1349-1350 1355-1356 1358 1364 1369 1379 1383-1384 1431 1435 1476 1507 1519 1532 1536 1547 1554 1564 1567 1578 1582 1587 1593 1595 1601 1608 1615 1619-1621 1638 1644 1661 1665-1666 1673 1687-1688 1690 1715 1723 1728 1749 1753 1757 1759-1761 1765 1771 1774 1776 1778 1781-1782 1786
fetal heart	Invitrogen	FHR001	105 124 160 289 864 1036 1148 1229 1614 1616 1762 1785
fetal kidney	Clontech	FKD001	5-8 11 40 47 57 65-66 82 85 102 124 163 171 216 222 224 235-236

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			258 277 280-281 307 310 314 330 371 387 392 395 403 422-423 431 436 443 455 469 500 519 522 542 563 572-573 585 600 619 623 650 654 657-658 660 679 719 731 780 798 821 833 844 854-855 857 864 868 878 911 929 958 960 969 990 992 1007 1046 1087 1103 1129 1139 1285 1312 1331 1355 1369 1371 1376 1391 1422 1425-1426 1440-1441 1470 1543 1598 1601 1618 1631 1651 1654-1655 1669 1678-1679 1691-1692 1733 1785
fetal kidney	Clontech	FKD002	352 384 426-427 440 583 602 1060 1131 1324-1325 1636
fetal kidney	Invitrogen	FKD007	20-21 82 163 335 679 988-989 1000 1227 1230 1320 1554
fetal lung	Clontech	FLG001	35-36 94 323 371 398 426-427 445 473 549 560 604 616-617 626 631 649 651 719 746 786-787 832 842 849-850 864 894-895 1075 1178 1182 1200 1206 1309 1311 1345 1429 1493 1567 1576 1620 1686
fetal lung	Invitrogen	FLG003	9 15-16 29 41 47 68-69 83 88-89 102 124 137 152-153 165 196 224 229 231 249 254 256 267 291-292 300 325 333 344-345 352 373 376 379 384 408 426-427 430 432 467-468 475 483 488 493 516 531 535 545 547 549 564 582 602 623 644 660 662-664 670 673 725-726 728 761 766-767 774 805 830 852-853 864 875 921 932 937 946 949 963 988-989 1014 1016-1017 1024 1027 1090 1097 1170 1185 1200 1215-1216 1224 1258 1290 1309 1320 1342 1347 1355 1369 1381 1413-1414 1431 1438 1449 1491 1512 1536 1547 1557-1560 1567 1590 1601 1636 1644 1653-1655 1662 1667 1671 1675 1680-1681 1706 1739 1760-1761 1769
fetal lung	Clontech	FLG004	103 276 334 465-466 737 843 1131 1614 1658
fetal liver-spleen	Columbia University	FLS001	3-11 13 15-21 25 30-39 41-48 50-51 54 56-58 60-66 68-69 72 75 77-80 82-83 85 87 89 92-103 105-110 112 116-124 126-127 130 133 135-139 141 144 147-149 152-153 157 163-165 167-172 174 176-178 180 186 188-190 193-194 196 198-200 202-206 210-214 219 221-231 233-236 240-244 246-247 250-251 255-256 258 261-265 268-269 272 274 276-278 280-281 284-286 288 293 295 299-301 304 306-307 309 311 314 316 318 320-321 326 329-332 342 344-345 350 352-353 356-358 360 362 370-374 376 378-384 386-387 390 392-393 400-401 403 406 408 410-412 415 417 419 422-437 439-442 444-445 448 452-454 456 459 461-470 472-479 481-483 487-488 490-491 493 500-501 503-506 509-513 515-520 522-524 526-529 531 534 536-540 542 547-549 553-554 561-562 564 567-568 571-576 579 581 583 585-597 599-605

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			607 610-613 615-621 623-624 626 628-634 636-640 644 647-650 655-660 665 669-670 672 674-675 678 681-682 684 690-695 697 702 708-710 713-714 716-719 725-728 730-731 734 736 738 740-741 743-746 748 750-751 759-766 768 772 774-777 779 783-788 793 796 798 800-805 808 810-812 814 818-819 821-824 826-832 834-837 843-847 849-867 869-876 878-883 887 889-895 897-898 902 904-914 916 919 921-928 930-937 939 945-950 953-958 960-961 963-965 967 969 971 974-978 980-983 986 988-990 992-993 995-997 1000-1002 1004-1008 1012 1014 1016-1019 1025-1026 1028-1031 1033 1035-1036 1039-1044 1047 1049-1050 1053-1056 1058-1059 1061-1064 1067-1070 1072-1074 1076 1078 1082 1085-1087 1089-1090 1097 1099-1103 1107-1113 1115-1119 1121-1123 1125 1127-1128 1131-1134 1136-1137 1144-1150 1153 1159-1160 1163 1170 1175 1177-1178 1188 1190-1192 1195-1200 1202 1206 1208-1211 1214 1216 1218 1221-1222 1225 1227 1234 1237 1241 1244 1246-1247 1251 1254 1258 1261 1266 1268 1270-1273 1277-1282 1284-1285 1287-1290 1294 1299-1300 1306-1308 1313-1320 1324-1325 1327 1330 1332-1333 1338 1341 1343 1345-1347 1349-1350 1353-1360 1362-1363 1365-1367 1369-1370 1372-1374 1376 1378-1381 1383-1384 1386 1389-1391 1400 1402-1403 1405-1410 1413 1415 1417-1419 1422-1429 1431 1435-1437 1439-1442 1445-1446 1448-1449 1454 1458-1459 1466-1470 1472 1474 1477-1478 1480 1482 1485 1491-1493 1496-1498 1501-1507 1509 1511-1512 1516-1519 1524-1526 1529 1532 1536-1541 1546-1547 1549-1550 1552-1554 1562 1564 1569 1572 1574-1575 1578 1581 1583 1587-1588 1591-1592 1594-1595 1597-1598 1600-1604 1611-1612 1614-1615 1617-1618 1620-1622 1624-1625 1627-1628 1630-1632 1634-1639 1645-1651 1653-1662 1664 1667-1669 1671 1673-1674 1676-1688 1690 1696 1701-1703 1706-1709 1711 1713-1714 1718-1719 1722 1724-1727 1731-1733 1738 1740-1741 1743-1744 1746 1748 1751-1752 1754 1760-1765 1767-1773 1780 1783-1786
fetal liver-spleen	Columbia University	PLS002	3-11 13 15-21 26 29 32 35-39 42 44-45 48 50-51 54-55 57-58 61 64 68-69 73-75 78 80 82 84 87 95-98 100 103 105 107-108 110 112-113 116-119 122-125 128 130 137-138 145 147-153 155 157 159 161-163 166 168 171-172 174-175 177 181 188-189 193-194 196-198 200-203

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			206 212-215 219-221 223 225-229 231-232 240-244 246-247 250-251 258-259 262 264 268-269 272 275 277 280-281 284 286 288 290-292 295 298-299 301-304 306 308-310 318 320-321 323 325 329 331 334 342 348-349 352-353 356 359 368 371 374 376-379 381-384 386-387 392-393 397-398 400-401 403 410-413 421 423 426-427 429-430 433-436 438 440 443 445 448 451-452 454-455 460-463 465-467 469 471-473 475-476 478-479 481-483 487 490-491 493-494 497 500-501 503-505 509-513 515-517 519-520 524 526-531 534 537-542 544 547 552-554 556 558 561-562 564-567 571-577 583-587 590-591 593 595 597 601 604-606 608-613 616-617 619-624 626-632 634 637-642 644 647 649-652 654-659 662-665 669-672 674-675 681-682 685 688 690 696 698 700-703 707 709-710 713 717 719-721 723-724 728 731-732 734 737-738 742-745 748 752 754 759 763-766 768 770 773-777 780 782 784 786 791 795-798 801-802 805 808 811-812 818 823-824 826-827 832 834-837 839 843 846 848-856 858-861 865 867 869 871 873-874 876 878 881-882 887 889 892 894-898 901-902 904 906-908 913-915 919 921-924 926-932 934-935 937 939-941 943 946-947 950 953 958 961 965-967 971 973-975 977-979 981 984-985 990 992-993 995-997 999 1001 1004-1007 1009-1011 1013 1016 1020 1023 1025 1027-1031 1033-1035 1039-1042 1044-1045 1049 1053 1055-1056 1058-1059 1062 1064-1065 1067-1070 1072-1074 1079 1082 1087 1089 1093 1097 1099-1103 1105-1107 1109-1114 1123 1125-1127 1132-1134 1140 1143-1145 1148-1150 1156 1158 1160 1163 1172-1173 1177-1178 1181-1184 1190-1192 1195-1197 1199 1204 1206 1208 1211 1214 1216 1219 1227 1230 1234-1235 1237 1240-1241 1243 1245 1247 1256 1258 1260-1261 1264 1268 1270-1271 1275 1278-1279 1284-1286 1288-1289 1299-1301 1306 1308 1312 1314 1317-1319 1323-1325 1327-1330 1334-1335 1339 1343-1347 1349-1350 1354-1355 1357 1360 1362-1363 1365-1367 1369 1372 1376 1378-1380 1386 1389-1391 1394 1400 1403 1406 1409 1416-1419 1422-1427 1429 1435 1437-1438 1440-1442 1446 1448-1450 1453 1460-1461 1468 1470 1472 1474-1475 1478 1482 1486 1490-1493 1496 1498 1500-1504 1506 1508-1509 1511-1512 1516 1518-1519 1521 1524-1528 1531 1536-1538 1543 1547 1550 1554 1556 1564 1567-1569 1580 1587-1588 1591-1592

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1597-1598 1600-1601 1611-1612 1618-1628 1630-1631 1635-1638 1641 1646-1649 1652 1654-1659 1661-1662 1664 1667-1669 1674 1676-1679 1683-1684 1686-1688 1691-1692 1699 1702 1707 1711 1713-1714 1717 1719 1722 1726-1727 1730-1733 1738 1740 1743-1744 1748-1752 1758 1760-1761 1763-1764 1767 1769 1772-1773 1776 1779 1783-1786
fetal liver-spleen	Columbia University	FLS003	103 300 318 321 352 372 379 381 384 392-393 403 422 424 429 434-435 440 444 453 503 515 544 592 978 1064 1324-1325 1327 1333 1357 1369 1378 1418 1424 1622 1646 1649 1680-1681 1689-1690 1717 1743-1744 1769
fetal liver	Invitrogen	FLV001	15-16 26 34 58 61 64 70 75 78 89 98 105 112 116 120-121 123 133 151 166 176 180 194-196 198 200 204-206 210-211 220 225-226 230 235-236 239 247 259 261 267 272 277 280-281 303 310 313 317 320-321 329 344 356 371 374 376 379-382 395 408 412 414 419 429 434-435 441-442 465-466 490 494 504-506 509 522 527 534 552-553 562 567 569-570 572-574 607 631 657-658 667 669 672 685-686 702 717 725-726 732 748 759 761 778 784 786 809 817 829 837 857 861 872-873 875 881 889 894-895 909 911 916 954 963 967 974 977 986 988-989 993 995 997 1000 1005-1006 1008 1014-1015 1020 1042-1043 1070 1086-1087 1089-1090 1118-1119 1122 1144-1145 1148 1153 1157 1159 1183 1195-1196 1227 1250 1257-1258 1262 1267 1280 1285 1307 1312 1314 1317-1320 1344-1345 1349-1350 1355 1362-1363 1403 1405 1415 1419 1425-1426 1429 1431 1442 1448 1463-1464 1469-1470 1489 1528 1536 1539 1549-1550 1557-1562 1577 1583 1598 1601 1611 1615 1622 1644 1649 1666 1674 1706 1721 1738 1746 1763-1765 1774 1776 1779
fetal liver	Clontech	FLV002	676 998 1719
fetal liver	Clontech	FLV004	93 133 214 301 355 374 379 555 581 601 679 837 847 859 1123 1236 1270 1313 1324-1325 1327 1355 1367 1425-1426 1536 1690 1733 1760-1761
fetal muscle	Invitrogen	FMS001	26 37-39 50-51 58 84 86 89 98 113 128 131-132 139 155 172 186 194 198 201 206 211 230-231 256 261 276 282 286 302 325 359 361 375 379 383 398 412-413 419 430 436 448 452 462-463 473 477 503 519 529 561 569-570 590-591 597 607 623 626 635 647 660 672 715 725-726 730 733 761 775-777 788 826 837 860 874 913 915 921 935 970 980 986 988-990 992 1000-1001 1007 1014 1027 1035-1036 1045 1060 1064 1070 1083 1097

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1099-1102 1116-1117 1121 1164 1173 1198 1208 1228 1240 1258 1266 1270 1277 1298 1317-1320 1324-1325 1329 1336-1337 1369 1383-1384 1399-1400 1403 1409 1433 1505 1514 1542 1551 1554 1557-1559 1562 1589 1599 1620 1632 1644 1650 1652 1671 1675 1712 1725-1726 1743-1744 1754 1766
fetal muscle	Invitrogen	PMS002	119 221 273 402 426-427 463 547 599 736 869 1000 1033 1083 1266 1431 1440-1441 1468 1545 1599 1673 1678-1679 1687-1688 1710 1712-1714 1723 1725 1731-1733 1743-1744 1760-1761 1767
fetal skin	Invitrogen	FSK001	1 4-11 15-16 20-23 25 29 33 40 43 46 56-57 60-61 64-66 75 82 87 97-98 105 107-108 113 118-119 123 133 135-137 139 144 146 148 151-153 156 163 170 176 180 188-189 197-198 200 202-203 210 218 222 231 246-247 261 263 265-270 277 285-286 290 293 299 301 307 311 321 325 328 330 333-335 339 341 345 351-352 355-356 358-359 362 368 370 372 376 379-382 384 388 394 404-405 408-409 411-412 419-420 424 426-427 436 441-442 445 448-449 454 462 465-466 472 476 490 493 504 506 509 515-517 519 526 531 537-540 547 549 560-561 567 572-573 581 584 589 611-612 615 623 630-631 635 647 649 651 657-658 660 662-665 667 669 672 676 678 681 688 701 704-705 709-710 713 717 720-721 725-726 728-729 732 748 750 753 759 764 766 770 775-777 780-781 786 788-789 798 809 811 814 816-817 822 824-826 831 842 857 859 861 863-864 881 894-895 908 910-911 916 918 922-923 928 932-933 935 937 946 948-949 953 960-961 966-967 970 975 977 986 990 992-993 999-1000 1004 1007 1013 1018 1025 1027 1032 1035 1041-1043 1054 1057-1058 1060 1062-1064 1069 1072 1077 1090-1091 1097 1099-1103 1108 1113 1119 1123 1128 1131 1134 1140 1148-1149 1152-1153 1156 1163 1167 1178 1182 1189 1192 1195-1196 1198 1201-1205 1208 1211-1212 1216 1219-1220 1222 1225 1240 1243 1258 1266-1267 1274 1277 1280 1282-1285 1299 1310 1317-1322 1324-1325 1329-1330 1342 1344 1346 1349-1351 1354-1357 1365-1366 1369 1371 1373 1376 1378 1380 1383-1384 1387 1399-1400 1405 1410 1427 1429 1431 1433-1435 1439-1441 1448-1449 1454 1457 1468 1470 1472 1475 1480-1481 1487 1490-1491 1493 1498 1509 1512 1521 1525-1526 1529 1535-1536 1547 1549 1557-1559 1588 1592 1595 1597-1598 1601 1603-1604 1608 1611 1614 1618 1624-

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1626 1632 1634 1636 1641 1643-1644 1646 1654-1657 1660-1662 1665 1668 1675 1685 1687-1689 1702-1703 1709-1710 1716 1719 1724 1727 1731-1732 1737-1740 1742 1747 1749 1755 1760-1761 1765 1772 1776-1777 1779-1780 1786
fetal skin	Invitrogen	FSK002	13 286 302 307 313 321 330 335 339 341 354 370 372 385 400 402 408 414 426-427 433 436 450 454 515 544 585 598 767 810 845 939 1076 1109 1155 1317-1320 1326 1333-1335 1343 1347 1350 1369-1371 1377-1378 1391 1397 1422 1466 1647 1656 1678-1679 1687-1688 1693 1718 1721 1725 1731-1732 1739 1755
fetal spleen	BioChain	FSP001	110 137 211 353 589 927 1108 1639 1771
umbilical cord	BioChain	FUC001	4-8 10 12 14 17 33-36 44-46 57 64 68-69 75 82 85 101 104 113-114 116 119 122-124 133 137 153-154 157 161 163 166-167 175 181-184 186 192 197-198 200-202 212-215 230 234 246-247 251 256 263 267 271-272 280-281 284 295 301 314 317 321 326 333-335 345 351 356 368 371-373 379-380 386 390 392 394 406 408-410 412 414 416 420 424 427 430-436 438 444-446 454 459 461 463 467 473 482-483 486 488 490 495 504 509 524 526 537-540 547 555 561 574-577 588-591 593 606 615 620-621 632 637 645-647 650 659-660 662-664 667-668 674-675 684 687 696 698 701 703-705 709 711 714 719-720 725-727 732 749-750 762 765 771 775-777 780 789-791 793 796 802-803 814-817 822 833 843 845 848 858 861 864 875 879 888 894-895 897-900 903 906-907 911-912 925 930-933 936 940 948 953 960 966 977 984 990 992 998 1000-1001 1005-1007 1016 1023 1025 1037 1046-1047 1059 1061-1063 1073 1076-1077 1089 1094-1097 1112-1113 1115 1134 1144-1148 1151 1154 1156 1163 1171 1197 1204-1205 1208 1216 1218 1224 1234-1235 1243-1244 1246 1279 1283 1286-1287 1298 1316 1320 1344 1346 1350 1357 1359 1371 1373 1375 1381 1398 1400 1403 1408 1414 1424 1427-1428 1431 1433 1440-1442 1446 1454-1455 1479 1482 1484-1485 1489 1492-1493 1504-1505 1513 1525 1527 1536 1538 1546 1565 1567 1571 1573 1575-1576 1578-1579 1591 1595 1600-1601 1608 1612 1615 1621 1624 1626 1636-1637 1647-1648 1651 1653 1656 1658 1661-1662 1672 1675 1682 1684 1686-1688 1690 1709-1710 1722 1727 1729 1735-1738 1740-1741 1760-1761 1768
fetal brain	GIBCO	HFB001	4 9 11-13 17-18 22-23 25 37-39 42-47 50-51 54-55 58 60-61 65-66

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			72 75 77 80 82 85 90-91 94 100-102 107 110 112-116 118-119 122-123 126 128 134 136-140 147-148 153-155 157 161 165 169-172 175 181 186 188-189 197-198 204-206 208 210 215 222-223 225-226 230 235-238 240-241 247 253 256-258 260-262 267-269 276 279-281 284 286 289 298 300-302 307 310 318 321-323 325 330-331 339 341 346-349 352 354 356-359 362 364-365 371-372 377 379-380 382 384 387 390 400 408 414-416 419 424 431 434-435 438 441-443 449 451 453-455 457-463 470 472-473 475 477-478 482-483 486-488 490-491 493 496 499-500 502-504 506-507 509-512 516 519-520 522 525-526 529-530 537-540 543-544 546-547 566-567 569-570 572-582 585 588 590-591 593 595 599 601 604 606-609 611-612 614-620 622-624 630 632 636 643 645-647 650-652 654 659 661 665 667-668 670-672 676 678 681 687 689 692-694 697 699 710 714 717 721 727 729-732 734 736 738 743-746 750-751 759 763 766 770 772 775-777 784 789 791 796 799 802-805 810-811 814 819-821 824 826 830 834-837 839-850 854-856 858-860 862 864 869 871 876-877 879 883 886-887 890-891 893-895 898-901 905 908-910 912-916 919 922-923 925 927 930-933 935-938 948 952-960 963-964 967 969-972 975 978-979 981 983 986-987 990 992 995 997 999-1002 1005-1009 1011-1013 1016 1018-1019 1023 1026 1029-1031 1033-1035 1038 1041 1047 1050 1053 1057 1059 1064 1068 1070 1072-1073 1078-1079 1081-1082 1086 1089 1094 1097 1103 1107-1109 1113-1115 1121-1122 1127 1134-1135 1138 1140 1143 1148-1151 1153 1156-1157 1159 1167 1170 1175 1193-1194 1200 1202 1207-1209 1211 1216 1219-1220 1226-1227 1229 1232-1234 1240-1241 1243 1246 1249-1251 1253-1254 1258 1267-1268 1271 1276 1279 1282 1285-1289 1293-1294 1305 1307-1308 1312 1316 1320 1327 1338-1339 1341-1344 1346 1349 1355-1357 1359 1365-1366 1369-1370 1373-1375 1379 1386 1389 1394 1398 1409 1413-1414 1416-1417 1420-1421 1425-1427 1430 1433 1437 1439 1442 1445-1452 1454-1457 1459 1463-1464 1468 1470 1474 1477-1479 1489 1492 1494 1497-1498 1501-1503 1507 1509 1511-1513 1517 1520-1521 1524-1526 1531-1533 1535 1537-1538 1547 1554 1556-1559 1564-1567 1571 1584 1587 1589 1594 1599-1601 1611-1612 1614-1616 1619-1620 1625-1628 1630-1631 1634 1637-1638 1640-1643 1645 1648-

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1649 1651 1653-1655 1657-1658 1664-1665 1667 1669 1673 1678-1679 1683-1684 1686 1693 1701 1704-1705 1709 1713-1714 1717-1720 1724 1727-1728 1731-1733 1737-1738 1743-1744 1752 1754-1755 1757 1760-1761 1765 1772 1779 1785
macrophage	Invitrogen	HMP001	5-8 110 204-205 503 634 678 859 878 933 988-989 1379 1448 1504
infant brain	Columbia University	IB2002	10 12-13 15-18 22-23 25 29 34 37-39 43 47 50-51 54-56 58 60-63 65-66 68-69 72-74 80 82-83 86 88-92 97 100 102-104 106-108 110 112-113 115-116 118 123 128 130 134-136 138-139 143 147-149 151-152 154-155 163 165-167 169 172-175 181-184 186 193-196 198 201 203-205 209-210 214-215 222 224-226 231-232 235-236 239 246-247 252 257 260 268-269 272 276-277 279-281 286 288 291-292 295 298 300-301 304 307 310 313 321-323 330-331 333-334 339 346-347 349 352 356-357 362 371-372 377 379-380 383-384 392 397 401 406 408 411 413-414 416 418-419 422 428 430-431 434-435 438 443 449 453-454 461 464-466 469-470 472-473 475-476 478 482-483 487 490 492 494 497 503 507-508 510-513 516 519-520 524-526 530-534 536-540 547 550-551 561 563-564 566-567 572-576 579 581-582 584-587 590-591 593 595-597 607-609 611-613 616-617 620 622-624 627 631 637 641 645-647 650-655 657-658 660-665 667-675 689 691 695 697 699 703 707 713-715 717 721 728-731 733-736 739 743 745 751 755 759 763 769-770 772 778 780-781 785 788-789 793-794 799 803 808 811 814 825-826 830 834-836 840-843 845 848-850 854-855 860 862 864-865 870 872 875-876 878 886 888 890-891 894-896 898 903-904 916-917 919 922-925 927-928 930-932 934-936 938 941 945-946 948-950 953-954 959-962 966-969 977 979 981 986-990 992 997 999-1000 1004-1006 1014 1016 1018-1019 1024-1025 1033 1036 1047 1051-1052 1054-1055 1057-1059 1063-1064 1068-1070 1073 1081-1082 1085 1089 1108-1113 1118-1120 1123-1124 1130 1132-1138 1140 1149 1151 1153-1154 1163-1170 1172 1174-1175 1183-1184 1188 1190 1193-1194 1196-1197 1199 1204 1208-1209 1211 1218-1222 1226-1227 1229 1231 1234 1241 1247 1249 1251 1256 1258 1261-1262 1269 1274 1279 1281 1283 1285 1287-1289 1294-1295 1305 1307 1313-1314 1316-1320 1329 1332 1341-1342 1345 1349 1356 1362-1363 1365-1366 1368-1370 1374 1381 1383-1384 1388 1400 1403 1406-1407 1413 1417 1420

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1423 1429-1431 1435-1436 1439-1441 1443 1447-1449 1451-1452 1454-1455 1457 1459 1463-1465 1468 1470-1471 1475 1479 1482-1483 1485 1493-1494 1496 1498-1499 1502-1503 1505-1507 1509 1522-1523 1525 1528 1531-1533 1542 1546-1547 1549-1550 1554-1555 1563 1565-1567 1569 1575 1580 1583-1586 1588 1590 1592-1593 1595 1598 1600-1601 1608-1610 1612 1614-1616 1619 1621 1624 1626-1627 1630-1633 1637 1639-1640 1642 1644 1647 1652 1654-1655 1658-1659 1664-1665 1672-1673 1676-1681 1685-1688 1693-1695 1701-1702 1704 1708 1717-1720 1723-1724 1726-1728 1733 1735-1741 1743-1744 1752 1755-1758 1762 1765 1771 1774 1777-1778 1786
infant brain	Columbia University	IB2003	17-18 20-23 29 34 43 60 68-69 78-80 88 100-101 107 110 112 118 123 128 133 135-137 146 148 152 159 166 169 174 194 198 203 215 223 225-226 229 235-236 247 260 276-281 286 290-292 295 300-301 310 322 324 331 334 339 346-347 349-350 352 357 371 376-377 382 384 403 408-409 414-415 453-455 472 476 478-479 490 503 507 516 520 530 534 536-540 551 563 572-576 585 587 590-591 593 595-596 601 606 612 616-617 620 622-624 650 652-653 661 665 670-671 674-675 678 689 715 717 727-728 730 734 759 775-777 780-781 785 796 806-807 811 824 845-846 864 869 875 882 889 894-895 898 904 917 919 921-923 932 935-936 946 950 954 962 977 979 997 999-1000 1005-1006 1009 1011 1017 1024 1033 1037 1043 1055 1057 1109 1114-1115 1120 1123 1127 1144-1145 1149 1151-1153 1160 1167 1170 1174 1193-1194 1196 1199 1202 1206 1209 1220-1221 1226 1229 1240-1241 1251 1258 1284 1288-1289 1305 1314 1327 1333 1344 1347 1350 1356-1357 1365-1366 1378-1379 1388 1400 1403 1421 1423 1431 1436 1440-1441 1446-1447 1457 1459 1471 1499 1503 1507 1509 1536 1546 1557-1559 1567 1572 1587 1595 1598 1610-1612 1615 1631 1639 1644 1647 1657-1658 1673 1678-1681 1683-1684 1701-1702 1708-1709 1713-1714 1719 1757 1760-1761 1765 1771 1778
infant brain	Columbia University	IBM002	101 113 139 152 260 279 290-292 374 377 551 563 608-609 653 659 814 954 1005-1006 1029-1030 1130 1164 1209 1258 1294 1305 1320 1327 1397 1431 1498 1507 1615 1640 1694-1695 1763-1764 1767 1779
infant brain	Columbia University	IBS001	10 12 119 175 279-281 321 334 371 446 551 563 623 652 667 669

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			671-672 819 949 966 1113 1130 1151 1188 1193-1194 1196 1229 1258 1265 1271 1287 1317-1319 1324-1325 1342 1423 1440-1441 1448 1471 1482 1525 1532 1546 1562 1569 1588 1591 1610 1618 1647 1649 1658
lung, fibroblast	Strategene	LFB001	5-9 17 20-21 25 68-69 82 94 105 153 157 197-198 203 207-208 212-213 223 262 266 283 302 321 326 333 356 370 427 430 436 446 462 472 493 498 503 516 519 527 535 537-540 542-544 562 565 567 586 599-600 607 615 630 647 662-664 692-694 712 719 745 748 775-777 794-796 810 837 843-847 849 854-856 869 876 903 934 953 955-956 964 975-976 984 1000 1005-1007 1024-1025 1033 1039 1053 1064 1070 1072 1082 1112-1113 1134 1136-1138 1140 1195 1223 1232-1233 1246 1279 1285 1295 1311 1320 1334-1335 1343 1427-1428 1446 1478 1482 1493 1504 1537 1552 1555 1567 1575 1582 1598 1620 1625 1632 1638 1645 1654-1655 1662 1680-1681 1684 1686 1690 1696 1702 1711 1733 1741 1760-1761 1778 1785
lung tumor	Invitrogen	LGT002	5-10 18 20-21 29 33-36 40 43 52 54-55 61 65-66 68-70 73-75 80 85 88-89 93-94 100 103 106-108 112-113 115-116 118-119 123-124 126 130-132 135-137 139-141 143-144 147-148 151-153 155-156 159 161 164 169 171 179-180 185 190 193 194 196-199 203-208 210 212-214 216-217 219 222 233 240-241 244 246 251-252 255-256 261-262 266 272 276-277 279-281 284 286 288 290 295 298 301-302 309-312 317 321 329 332 341-342 344-345 348 352 358-360 363 368 370-371 376 380-381 384 389-390 398 400 409 414 423 426-427 430 432-436 443-444 450-451 454 462 468 472-477 480-483 487-488 490-491 493 496-498 500 503-506 509-512 515-516 519 521-523 526 530 534 541 544 547 554 557 564 566-567 572-576 585-586 588-589 595-596 601 607 611-612 615 619 621 623 626 630 632-633 644 647 649 651 655-656 660 662-665 667 669 672 683-684 696 700 706 710 713 716 718-719 722-723 728 734-739 743 750 752 763 765-766 773-778 784-785 787-789 791 800 802-803 809-812 814 824 826 828-829 832 838-839 841-845 849-850 852-855 857-861 864 866 874 878-880 882 887 890-891 897-898 902 904 906-907 910 916 918-920 922 924-925 927 930-932 934-935 937 947 950 953 955-956 961 963 966-967 969 971 977-979 981 984 986-987 990 992-993 995 997 999-1001 1005-1007 1009 1012-1013 1018 1020 1022-1024 1026 1029-1030 1033 1038 1041

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1045 1047-1050 1052 1054-1055 1059 1063-1064 1067-1071 1073-1074 1078 1085 1087 1089 1095-1097 1104 1106-1107 1109 1112 1116-1117 1119 1126 1134-1135 1139 1141-1142 1144-1145 1148 1152-1153 1156-1158 1167 1170 1172 1178 1195-1196 1198-1200 1202 1204 1208 1214 1216 1219 1222 1227 1234 1241 1247 1252 1257-1258 1265 1267-1270 1276 1278 1280-1281 1283 1285 1288-1289 1295 1300 1305 1308 1312 1317-1321 1329 1338-1339 1341 1344-1346 1349-1351 1353-1355 1357 1365-1366 1369 1378-1379 1383-1385 1394 1397 1400 1402-1403 1408 1417 1419 1423-1426 1431 1433-1436 1438 1444 1446-1448 1454-1455 1460 1466 1468 1470 1474 1480-1481 1483 1486-1488 1490-1491 1494-1496 1506 1508-1509 1511-1512 1515-1516 1519 1523-1524 1528-1529 1536-1540 1546 1549-1550 1555 1560-1561 1565 1567 1569 1575 1588 1591 1593-1594 1596-1598 1600-1602 1608 1614-1616 1618 1620 1624-1625 1627-1632 1636 1639 1644-1645 1647-1649 1652-1653 1656-1662 1664 1666-1667 1670-1671 1673-1675 1678-1679 1683 1685-1688 1690-1692 1696-1699 1705 1709 1716-1717 1722 1727 1730 1735 1739 1741 1743-1744 1748-1749 1753 1760-1762 1765 1767 1770-1771 1773 1775-1776 1778-1779 1786
lymphocytes	ATCC	LPC001	4 11-12 18 24-25 30-31 48 50-51 56-57 68-69 80 92 98 103 105 110 126 137 152-153 157 165 172 188-189 197 203 210 217-218 222-223 225-226 229 231 247 251 256 264 272 280-281 284 300-301 321 325-326 339 348 352 357 371 382 384 390 400 404 412 414 421 423 426-427 430-431 445 447-448 451 454-455 475 503 516 526-527 530 537-540 549 556-560 563 574 577 589 602 613 615-617 621 623 628-630 636-637 647 649 657-659 690 697 717 723 755 764 775-777 780 786 789-790 793 800 802 822 838 849 866 869 876 881-883 892 898 906-907 911 921-923 928 975 990 992 996 1001 1004-1007 1033 1050 1054 1078 1107 1135 1140-1141 1143 1148 1158 1163 1177 1199 1205 1216 1226 1231 1236 1241 1244 1250 1258 1260 1265 1269-1271 1290-1293 1308 1312 1317 1319-1320 1339 1345-1346 1348 1350-1351 1357 1367 1369 1379 1381 1383-1384 1386-1387 1389 1394 1397 1405 1423 1425-1428 1431 1437 1446 1448 1461 1466 1470 1472 1474 1482 1492 1506 1528 1537 1546 1549 1591 1598 1600 1603-1604 1606 1627 1636

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1638 1647-1649 1651 1658-1659 1664 1676-1677 1680-1681 1687-1688 1699 1711 1715-1716 1726 1728 1737 1740 1746 1748 1752 1756 1758 1777 1779
leukocyte	GIBCO	LUC001	3-4 10-11 13 15-18 20-21 24-25 30-31 35-36 40 43-45 48 50-51 54-58 60-63 68-69 75 79-80 82-83 85 88-91 93-96 98 100 103-104 107-108 112 116 119 123 125-128 134-140 142 147-149 151 153 155 157 162-163 167 169-172 174 177-179 186 190 192-199 203-207 210 212-215 217-219 222-223 229 235-236 247 251 255-258 260 262 272 274-277 280-281 285-286 297-301 307-310 313-314 316-317 321 325-330 333-334 340-342 348-349 352 354-358 370-371 380-385 387-388 400 405 408-410 412 414-416 421-425 430-431 434-435 437 439 441-442 445-451 453-454 456 459 461-464 468-472 474-479 481 483-485 487-491 496 499-501 503-504 509-513 516-519 522 526-527 529-531 534 536-540 542 547-549 553-559 566-567 571 574-577 579 582 584-586 589 593 595-597 601-602 604 606-607 611-613 615-621 623 627-629 633 636-637 642 644-650 655 659-660 662-665 667 669 674-675 678 682-684 692-696 698 700 706 708 710 716-720 725-726 729-736 738-739 743-746 749 751 753 756 759 765-766 768 770-778 780 784-786 788-790 793 796 798 800 802-803 810-811 814 817 819 826 828-830 832 834-836 838 843 845-860 863-864 866-871 877-879 881-892 894-896 898 902 904-914 916 919-925 927 930-932 935-936 941-942 945 948-949 953 955-956 958 960-962 964 967 970-971 973 975 977 985-990 992-993 995-996 999-1002 1004-1009 1011 1014 1017-1019 1022-1023 1025 1027 1029-1031 1033-1036 1038 1041 1043 1047 1050 1053-1054 1058-1059 1061-1062 1064 1068 1070 1072 1078 1085-1086 1089-1091 1093 1097 1106-1107 1110-1113 1115-1117 1122-1123 1125 1129 1132-1133 1135-1137 1140-1145 1152 1158 1163 1168 1170-1174 1176-1178 1180 1182-1183 1186 1195 1198-1200 1202 1205-1206 1211 1216 1219-1221 1223-1227 1230-1236 1238-1242 1247 1252 1254 1256 1258 1261-1262 1264-1265 1269-1270 1272-1275 1277 1280-1284 1287-1293 1299-1300 1306 1308 1312-1313 1317-1320 1322 1324-1330 1333-1335 1339 1341 1343-1347 1349 1353-1357 1359-1361 1365-1367 1369-1370 1373-1374 1377 1379-1381 1386-1387 1394 1400 1403 1409 1419 1423 1425-1428 1430-1431 1433-1434 1437-1438 1440-1442 1446-1448 1450

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1453 1458-1459 1463-1464 1468 1470-1471 1474 1477-1478 1482-1488 1490-1493 1496-1501 1504 1506 1509 1512-1513 1516 1519 1521-1522 1524-1525 1527-1528 1531 1534 1538 1541 1545-1547 1549-1550 1553 1555-1556 1560 1565 1567 1575 1580 1589 1591 1594 1596 1598 1600-1602 1606-1608 1611 1614 1620-1621 1624 1626-1629 1631-1632 1636 1638-1639 1641 1644-1645 1648-1650 1653-1655 1658-1660 1662 1669-1670 1675-1679 1684-1688 1690-1692 1696 1700 1702 1707-1709 1711 1716-1717 1720 1723 1725-1727 1733 1737-1738 1741 1743-1744 1748-1749 1752 1755 1760-1762 1765 1769 1771-1772 1781-1784 1786
leukocyte	Clontech	LUC003	4 35-36 44-45 61 68-69 75 82 102 119 139 154 179 197 244 280-281 324 372 404 430-431 455 461 476-477 481 503 537-540 554 575-576 581 589 608-609 621-622 624 630 632 647 662-664 669 679 698 764 773 775-777 802 848 851 856-857 879 905-907 915 949 952 990 992 1002 1113 1119 1170 1183 1216 1236-1237 1241 1275 1346 1353 1357 1359 1377 1506 1515 1534 1553 1591 1600 1613-1614 1621 1628 1670 1676-1677 1691-1692 1699 1733 1738 1772
melanoma from cell line ATCC #CRL 1424	Clontech	MEL004	25 35-36 43 80 104 126 128 150 163 166 188-189 197 210 215 220 271 277 280-281 310 317 336-338 345 351 372 380-381 383 387 412 415-416 430 445 448 454 456 467 481 490 499 503 526 528 546 548 567 575-576 588 601 613 615 647 660 665 734-735 737 759 778 787 790 800 832 845 856 859 869 878 883 887 905 914 932 934 958 976 985 990 992 999-1000 1025 1031 1038 1050 1055 1068 1074 1088 1099-1102 1107 1136-1138 1149 1156 1163 1172 1190 1195 1200 1214-1215 1217 1226-1227 1235 1238-1239 1244 1253 1278 1280 1293 1311 1320 1330 1334-1335 1345 1355 1367 1386-1387 1394 1403 1406 1414 1423 1437 1442 1465 1521 1529 1536 1539 1541 1547-1548 1582 1620 1626 1631 1638 1647 1653 1660 1667 1669-1670 1680-1681 1696 1704 1715 1724-1725 1731-1732 1750 1760-1761
mammary gland	Invitrogen	MMG001	5-8 10 12 14-18 20-21 24-25 29 33-39 42-43 52 55-58 60-64 68-69 71 73-74 79-80 82 89 98 100 103 106 108 112 123 128 133-137 144-146 148 150-152 154 158-159 165-166 170-172 174 176 178 181-185 188-190 194-198 201-206 210 217-222 224 227-228 231 233-237 247 251 253-254 256 261-263 266-267 271 276-277 279-281 284-286 288

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			290 297 299 301 304 309-312 318 320-321 323-325 327-329 331-332 334 339 341 344-345 348 350 356 359-360 362-363 368 371 376 379-383 388 390 393-395 397-398 405 408 412 414-415 423 430 434-437 441-444 448 451-455 462-464 474 476 479 482 485-486 488 490 494-495 498 503 506 509-512 516-517 519-520 522 527 529 534 537-541 547 549 554 557 562 572-574 587 589-591 597 602 607 618 623 628-629 632 634-640 644 647-648 650-652 655 657-658 660 665 667 669-672 674-676 679 682 688 695-696 706-707 710 713 717 720 722-730 732-734 736 738 743 747-748 750 755 759 761 766 770 780 784 786-789 794 803 806-807 809 814 817-822 827-829 837 842 854-858 863-864 866 869-870 872 878 881 889 893-900 904 906-907 911 916 919 921-923 926 935-937 946 948-949 953-954 957 960-961 963 965-966 970 977-978 984-989 993-997 1000-1001 1005-1006 1008 1013-1014 1016-1017 1023 1025 1027 1032-1033 1036 1039 1043 1045 1055 1057-1058 1063 1068-1075 1077-1078 1085 1087 1089-1091 1095-1102 1107-1108 1112-1119 1121-1123 1131-1133 1136-1137 1139-1142 1144-1145 1148-1149 1153 1159 1167 1170 1172-1173 1183-1185 1190-1192 1196-1199 1207-1208 1212 1216-1218 1222-1223 1225 1231 1234 1240-1241 1247 1253-1254 1258-1259 1261-1262 1270-1280 1283 1285-1286 1298 1307 1314 1316-1320 1323-1325 1330 1334-1335 1342-1345 1349-1352 1354-1355 1359 1369-1370 1377 1379 1381 1383-1384 1389 1405 1414 1419 1421-1423 1425-1426 1428-1429 1431 1434-1437 1439 1448-1449 1454 1457 1460-1464 1466 1471 1480-1483 1487 1489-1491 1493 1505 1507 1512 1519 1526-1528 1532 1534 1536 1539 1542 1547 1549-1550 1554 1561-1562 1564 1567 1572 1576-1579 1581-1582 1587-1588 1592 1594 1596-1597 1601-1602 1607-1608 1610 1612-1616 1618 1621-1622 1625-1626 1631 1635-1636 1641 1643-1644 1647 1650 1652 1654-1655 1657-1658 1660 1662 1664-1666 1669-1671 1673-1674 1676-1677 1680-1685 1689-1692 1701 1706 1713-1715 1719-1720 1723-1728 1730-1732 1738 1740 1742-1744 1746-1747 1749 1751 1753 1760-1762 1765-1768 1771 1774 1776-1777 1779 1783-1784 1786
induced neuron cells	Strategene	NTD001	29 35-36 80 116 123 156 163 181 214 230 280-281 284-285 307 321 330 340 358 371 375 377 380 382 422 424 492 497 532-533 542 546

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			549 566 586 595 612 645-647 654 734 775-778 780 792 799 821 826 856 858 875 936 953 985 990 992 1041-1043 1055 1072 1104 1193-1194 1206 1223 1246 1253 1274 1288-1289 1291 1294 1311 1320 1349 1359 1412 1423 1485 1620 1623 1645 1684 1705 1715 1751
retinoid acid induced neuronal cells	Strategene	NTR001	5-8 78 268-269 277 383 431 506 623 677 731 999-1000 1199 1425-1426 1547
neuronal cells	Strategene	NTU001	29 65-66 80 82 110 119 146 152 166 174 181-185 198 227-228 253 284 309 325 332 334 336-338 375 391 393 406 414-416 454 465-466 470 488 503 506 510-512 519 537-540 572-574 597 602 607 623 647 661 700 702 716 743 771 792 858 904 948 954 977 1000 1005-1006 1025 1064 1068 1122 1148 1185 1219 1226 1234 1246 1271 1283 1295-1296 1311 1317-1320 1329-1330 1350 1355 1365-1366 1378 1383-1384 1400 1412 1445 1505 1539 1547 1578 1647 1656 1683 1690 1738 1749 1783-1784
pituitary gland	Clontech	PIT004	311 314 379 408 419 430 454 1055 1095-1096 1272-1273 1312 1320 1378 1652 1671 1720 1725 1736 1741 1755
placenta	Clontech	PLA003	5-8 124 208 277 370 843 906-907 1280 1317-1319 1369 1609 1621 1737
prostate	Clontech	PRT001	9 46 57 71 107 147 171 177 197 201 229 231 242-243 274 280-281 307 310 317 330 358 373 382-383 400 430 434-436 461-462 469 477 489 497 500 505-506 513 521 526 531-533 547 618 649 657-658 662-664 710 729 767 771 789 820 861 871 874 890-891 905 938 945 963-964 988-989 1002 1025 1033 1045 1061 1095-1096 1112 1125 1142 1196 1198 1202 1232-1233 1241 1258 1272-1273 1287 1295 1313 1333 1341 1344 1349 1360 1362-1363 1367 1437 1442 1447 1475 1478-1479 1482 1489 1513 1517 1527 1531 1536 1598-1599 1628 1636 1657 1680-1681 1687-1688 1717 1738 1743-1744
rectum	Invitrogen	REC001	17-18 29 33 62-63 71 73-74 83 86 113 126 146 153 158 167-169 195 200 206 261 309 312 341 344 368 373 388 395 408 414 420 430 441-442 446 448 464 468 483 517 537-540 547 567 585 589 602 623 628-629 632 645-647 651 657-658 669 717-719 721 725-726 738 748 750 756 762-763 766 770 774 790 819 825 843 849 851 881 903 909 948-949 960 986 996 1020 1023 1033-1034 1064 1067 1070 1075 1086 1108-1109 1113 1130 1139 1153 1159 1172 1178 1185 1187-1189 1205 1220 1225 1240 1244 1271 1317-1320 1323 1334-1335 1350-1351 1355 1369 1373 1375 1425-

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1426 1436 1439 1469 1471 1477 1482 1546 1587-1588 1592 1596 1610 1622 1627 1644 1658 1662 1665-1666 1669 1675-1677 1749 1786
salivary gland	Clontech	SAL001	10 55 97 103 110 140 149 152 158 198 217-218 242-243 256 301 308 312 321 333 351 354 360 410 437 448 473 487 494 496 501 535 555 569-570 572-573 590-591 624 636 651 759 762 764 768 771 788 800 809 826 848 865 879 906-907 925 933 963 1016 1020 1025 1040 1046 1055 1066 1103 1150 1172 1181 1234 1281-1282 1288-1289 1298 1315 1320 1333 1336-1337 1346 1359 1373 1379 1424 1447 1449 1474 1482 1492 1494 1498 1511 1523-1524 1537 1554 1596 1626-1627 1636 1652-1655 1658 1665 1671-1672 1691-1692
salivary gland	Clontech	SALs03	158 326 1423 1463-1464
skin fibroblast	ATCC	SFB001	1320 1400
skin fibroblast	ATCC	SFB002	262 736 1025 1253
skin fibroblast	ATCC	SFB003	709 1119 1350 1631 1653
small intestine	Clontech	SIN001	25 142 146-147 151 155 198 203 244 260 271 280-281 286 288 298 301-302 308 312 334 340 371 398 408 412 414 416 423 426-427 430 434-435 445 452 454 478 503 516 519 521 523 543 547 549 555 559 563 569-570 585 592 604 611 626 628-629 632 650 659 681 710 714 718 750 764 780 798 829 842 857 859 866 887 892 894-895 901 904 906-907 912 919 935 997-998 1000 1007-1008 1026-1028 1044 1055 1089 1097 1116-1117 1131 1148 1169 1199 1219 1234 1247 1264 1279 1316 1320 1326 1341 1343 1349 1351 1374 1387 1398 1400 1403 1407 1423 1428 1468 1498 1501 1521 1550 1556 1585 1597 1636 1638-1639 1645 1653 1656 1662 1671 1675 1684 1691-1692 1704 1711 1717 1719 1722 1725-1726 1729 1733-1734 1743-1744 1762 1767 1780 1785
skeletal muscle	Clontech	SKM001	18 20-21 82 84 101 118 134 148 151 153 166 225-226 258 274 277 289 329 361 412 414 424 440 452 459 470 488 503-504 537-540 647 660 673-675 715 773 780 786 830 905 922 950 963 982 990 992 1020 1047 1063 1115-1117 1121 1134 1228 1268 1284 1298 1321 1329 1336-1337 1343 1409 1413-1414 1509 1599 1624 1644 1653 1712
skeletal muscle	Clontech	SKM002	168 1683 1712
skeletal muscle	Clontech	SKMs03	235-236 1409
skeletal muscle	Clontech	SKMs04	235-236
spinal cord	Clontech	SPC001	4 9 11 17 30-31 35-36 43 46 60

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			82 85 92 94 108 110 116 139 157 167 198 204-205 210 215 229 256 259 277 280-281 300-302 304 315 317 372 379 387 392 419 426-427 430 433 448 467 473 487 489 506 509 513 519 524 526 537-540 543 547 549 551 559 567 569-570 593 607 616-617 623 625 637 649-650 652 657-658 670-671 673 679 681-682 709 711 715 719 728-729 734 749-750 753 775-777 781 789 791 809 820 832 834-836 847-849 854-855 858 861 864 871-872 875 884 898 906-908 917 919 924 934 942 944 970 985 990 992-993 998 1013 1039 1053 1059 1065 1072 1075 1077 1082 1085 1097 1103 1109 1116-1117 1128 1134 1151 1170 1174 1192-1194 1215 1225 1241 1243 1283 1294 1307 1312 1320 1323 1327 1330 1350 1353-1354 1356 1359 1368 1375 1400 1406-1407 1423 1429 1437 1443 1448 1454 1470 1482 1492 1501 1508 1511 1529 1538 1548-1549 1565 1571 1578 1598 1600 1614 1625 1627 1630 1639 1646 1651-1652 1670 1686 1696 1740 1751 1755 1771
adult spleen	Clontech	SPLc01	117 312 326 348 424 426-427 431 845 866 1320 1330 1333 1344 1355-1357 1371 1387 1397 1446 1538 1579 1669 1686 1739 1767
stomach	Clontech	STO001	10 15-16 61 68-69 100 117 149 197 201 227-228 231 249 273 280-281 287 291-292 302 312 358 362 426-427 430 446 462 475 479 535 597 620 630 651 662-664 722 739 780 782 785 846 919 960 964 966-967 976 1008 1012 1032 1042 1063 1071 1135 1170 1208 1234-1235 1259 1277 1280-1281 1322 1349 1359 1369 1449 1468 1474 1478 1487 1493 1498 1557-1559 1622 1634 1651 1653 1729
thalamus	Clontech	THA002	9 11 25 85 87 112 137 146 180 190 198 206 210 212-213 235-236 239 261 268-269 279 290 301 325 333-334 341 351 356 364-365 379 388 393 396 419-420 441-442 458 477 483 508 525 531 549 567 606 608-609 647 681 715 725-727 736 774 782 784 794 827 883 890-891 899-900 961 997 999-1001 1004 1034 1055 1097 1129 1144-1145 1150-1151 1157 1172-1173 1177 1193-1194 1208 1220 1249 1280 1305 1345 1355 1369 1434-1435 1440-1441 1454 1496 1546 1549 1562 1572 1578 1590 1594 1613-1614 1640 1651-1652 1671 1687-1688 1703 1743-1744 1746-1747 1753
thymus	Clontech	THM001	44-45 54 57-58 62-64 79 104 123 126 134 153 193 212-213 218 242-243 258 274 277 279 297 301 307 327 330 333 342 351 358 371 410 430 445 465-466 468 471 483 487 493 503 506 509 517 526 535 537-

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			540 546 548 554 567 584 586 590-591 604 612 621 638-640 645-647 649 656 660 665 670 698 710 720 728 735 739 746 759 762 766-767 775-777 780 784-785 800 802 809 824 826 828 845 851 858-859 864 866 870-871 878 884 887 892 899-900 927 930-931 967 983 986 990 992 999 1014 1029-1030 1033 1059 1066 1073 1103 1107 1113 1116-1117 1119 1140-1142 1158 1163 1172 1177 1195 1206 1209 1213 1216 1218-1219 1221-1222 1227 1271 1277 1282 1320 1329 1349 1367 1369 1383-1384 1417 1419 1423 1425-1427 1448 1477 1488 1493 1536 1554 1620 1644 1646 1549 1654-1655 1661-1662 1669-1670 1674 1676-1677 1685-1688 1707 1711 1731-1732 1737
thymus	Clontech	THMc02	5-9 15-21 25 33 35-36 43-45 48 50-51 54-55 60 75 83 87 89 93 98-100 102 105 112 117 135-137 141 143 146 157 167 169 192 196 211 217-219 222 224 229 233 235-236 240-241 244 251-252 256 261-262 268-269 286 288 290 295 297 301-302 309-310 315-317 321 324 327 334 342 350 352-353 360 370-373 382 384 400 403 410 414-416 424 430-431 436 445 454-456 461 464-467 470 472 474-476 483 488 497 500 504 506 513 516 519-520 524 526 530-531 534 537-540 549 554-555 565-566 569-570 572-573 575-577 586-587 595 603-604 606 612 630-632 634 636 647 650 657-660 666-667 669 673-675 678 698 700 703 708 720 725-726 731 738-739 743-744 750-753 757 759 763-765 767 772-779 787 789-790 798 800 810 823 829 834-836 841 848 854-856 859 861 864 870-871 881 890-891 898 908-909 913 928 933 941 949 958 961 963 967 969 975 981 986 988-990 992 999 1007-1008 1014 1016 1039 1041 1073-1074 1079 1089 1097 1109 1114-1117 1122 1131 1140-1141 1144-1145 1163 1172 1175-1177 1186 1196 1198 1206 1211 1216 1220 1223 1227 1234-1243 1261-1262 1267 1271 1280-1281 1284 1290 1308 1317-1320 1322 1324-1325 1327 1330 1334-1335 1339 1346 1350-1351 1355 1357 1360 1370 1374 1377-1379 1386 1389-1390 1392 1397 1400 1402 1406-1407 1417 1423 1425-1427 1440-1441 1466 1474 1477 1483 1493 1498 1504 1506 1525 1536 1545 1549 1566 1594 1598-1600 1608 1611 1614 1621 1623 1625 1632 1639 1641 1644 1647 1649 1653-1656 1658 1662-1663 1671 1673 1678-1681 1686-1688 1693 1705 1707 1711 1717-1718 1726-1727 1731-1733 1737-1738 1743-1745 1758-1761 1771-1772 1779 1786

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
thyroid gland	Clontech	THR001	4 9-10 20-21 37-39 48 50-51 54-57 60-61 65-66 71 83 94-96 98-100 102 104 110 112 115-117 119 123 127 133 136-137 140 149 152-153 155-158 163-164 168-169 171 186 190-192 197 201-203 219-220 229 233-237 246-247 253 256 258 262 265-266 268-269 277 280-281 284-286 288-289 298-299 302 309-311 317 321 326 332 335 341-342 344 348 350 354 358-359 363 368 371-373 382-383 385 394 398 400-401 411 414-415 421 424 430-431 433-436 443-446 450-452 454-455 458 472-474 476-478 482 484-485 487-488 490-494 496-497 500-501 503-504 506 509-513 516-517 519 524 526-527 529 535-540 547 549 562 564 569-570 575-576 588 594-595 601-602 604 606 610 612 615-617 619-623 628-630 634-635 642 647 649-651 660 662-665 668 670 681 690-694 696 698 700 709 721 727-729 732 734 738 740-741 743 745 750 759 761 763 765 770 773 780 785 795-796 798 802 804 823-824 826 828 833 838 841-845 847 849 857-860 867 874-875 878 880-881 887-888 890-892 894-895 898 908 910-911 913-914 922-923 926-927 929 932-934 937 939 941-942 948 953 957 961 963-964 966 978-979 981-982 987 990 992 1001 1004-1006 1010 1014 1020 1024 1033 1038-1039 1044 1047 1050 1052-1054 1056 1058 1068 1070-1071 1077-1079 1088 1094-1097 1105-1106 1112-1113 1116-1117 1124 1126 1128-1129 1131 1134 1136-1137 1142-1143 1146-1147 1149-1150 1156 1161-1164 1167 1170-1173 1177-1181 1190 1192 1197 1200 1204 1208-1209 1214 1217 1219 1222 1230 1232-1233 1235 1241 1245 1247 1254 1257-1258 1260 1262 1271-1273 1283 1286-1289 1299 1306 1314 1320 1330-1332 1334-1335 1342 1345 1349 1365-1367 1370-1372 1374 1381 1394 1407 1419 1428 1436-1437 1440-1441 1443 1446-1449 1454 1459 1461-1462 1468 1470-1471 1475 1477 1479 1482 1491 1497-1498 1504-1505 1507 1513 1522 1524-1526 1528 1531 1534 1536-1537 1548 1550 1553 1555-1559 1562 1567 1578 1590-1591 1597 1599-1601 1612 1614 1616 1619-1620 1622 1624-1626 1628 1631-1632 1634 1636 1639 1644-1645 1648 1651 1653-1656 1658 1660 1662-1663 1667 1669 1671 1675 1678-1681 1683-1686 1689 1691-1692 1703 1709-1711 1717 1724-1726 1729 1734 1737-1738 1740 1743-1744 1749 1753 1759-1761 1770 1777 1786
trachea	Clontech	TRC001	9 29-31 46 48 87 104 107 110 135 158 222 262 266 286 301 318 331

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			352 372 377 384 414 424 445-446 454 472 474 491 496 560 579 588 593 597 607 612 626 681 702 719 810 859 866 878 894-895 912 916 922 932 935 1046 1075 1080 1099-1102 1113 1208 1215 1232-1233 1237 1281 1312 1385 1387 1405 1414 1424 1430 1437 1447 1505 1569 1579, 1586 1600 1641 1653 1667 1671 1676-1677 1683 1691-1692 1711 1717 1726 1772
uterus	Clontech	UTR001	17 19 25 41 46 57-58 61 89 104 108 139 152 174 198 200-201 206 263-265 274 290 387 408 420 438 446 448 452 473 491 493 499 503 506 513 519 522 526 530 542-543 560 601 610 632 659 665 720 751 773 780 833 845 857 872 877 912 929 934 937 996 1009-1011 1018 1050 1075 1107 1124 1170 1219 1258 1279 1287 1310 1320 1323 1343-1344 1375 1437 1451-1452 1478 1481 1498 1519 1521 1536 1552 1579 1597 1602 1606 1620 1626-1627 1649 1652 1661 1670 1719 1722-1723

TRADOCS:1416191.1(%CQN01!.DOC)

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1	Y41736	Homo sapiens	Human PRO1114 protein sequence.	1398	100
2	Y66656	Homo sapiens	Membrane-bound protein PRO943.	2389	99
3	AF113136	Homo sapiens	IL-1 receptor-associated-kinase-M; IRAK-M	3043	100
4	AF017806	Mus musculus	Zn-15 transcription factor	6351	77
5	X02761	Homo sapiens	fibronectin precursor	10535	98
6	X02761	Homo sapiens	fibronectin precursor	8990	89
8	X02761	Homo sapiens	fibronectin precursor	12564	99
9	AJ011679	Homo sapiens	Rab6 GTPase activating protein, GAPCenA	5251	99
10	W88501	Homo sapiens	Human stomach carcinoma clone HP10415-encoded protein.	2381	100
11	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	11336	98
12	Z97630	Homo sapiens	dJ466N1.4 (novel protein similar to ANK3 (ankyrin 3, node of Ranvier (ankyrin G))	896	100
13	Y58620	Homo sapiens	Protein regulating gene expression PRG2-13.	1894	98
14	AF213457	Homo sapiens	triggering receptor expressed on myeloid cells 2	1238	100
16	AF233453	Homo sapiens	RACK-like protein PRKCBP1	3124	99
17	AF201303	Homo sapiens	dhfr oribeta-binding protein RIP60	3130	98
18	AF064205	Homo sapiens	dynactin 1 p150 isoform	6377	100
19	U00059	Saccharomyces cerevisiae	Yhr121wp	174	26
20	AB032903	Homo sapiens	guanosine monophosphate reductase isolog	1801	99
21	AB032903	Homo sapiens	guanosine monophosphate reductase isolog	1485	99
22	AF140507	Homo sapiens	Ca2+/calmodulin-dependent protein kinase kinase beta	3083	99
23	AF140507	Homo sapiens	Ca2+/calmodulin-dependent protein kinase kinase beta	2300	99
24	AJ289131	Homo sapiens	chondroitin 4-O-sulfotransferase	2211	99
25	U33460	Homo sapiens	DNA-directed RNA polymerase I, largest subunit	8777	98
26	Y44488	Homo sapiens	ACRP30R2 variant protein.	1387	100
27	U43701	Homo sapiens	ribosomal protein L23a	791	100
28	U02032	Homo sapiens	ribosomal protein L23a	767	97
29	Y41324	Homo sapiens	Human secreted protein encoded by gene 17 clone HNFIY77.	1083	99
30	W71749	Homo sapiens	Human ubiquitin conjugation system protein 2.	715	90
31	W71749	Homo sapiens	Human ubiquitin conjugation system protein 2.	631	82
32	AF231917	Homo sapiens	long-chain 2-hydroxy acid oxidase HAOX2	1811	100
33	Z29481	Homo sapiens	3-hydroxyanthranilic acid dioxygenase	1507	99
34	AB001451	Homo sapiens	Sck	2869	100
35	Y00644	Homo sapiens	precursor polypeptide (AA -34 to 287)	1667	99
36	Y00644	Homo sapiens	precursor polypeptide (AA -34 to 287)	1104	98
37	Y78795	Homo sapiens	Human antizuai-2 (AZ-2) amino acid sequence.	3586	78
38	Y78795	Homo sapiens	Human antizuai-2 (AZ-2) amino acid sequence.	4726	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
39	Y78795	Homo sapiens	Human antizuai-2 (AZ-2) amino acid sequence.	3556	77
40	U93121	Homo sapiens	M-phase phosphoprotein-1	3747	100
41	Y42750	Homo sapiens	Human calcium binding protein 1 (CaBP-1).	795	100
42	AF282626	Homo sapiens	latexin	1189	100
43	G02150	Homo sapiens	Human secreted protein, SEQ ID NO: 6231.	384	94
44	U19617	Mus musculus	Elf-1	2724	88
45	U19617	Mus musculus	Elf-1	2062	86
46	AF100758	Homo sapiens	osteoinductive factor OIF	1538	100
47	Y87591	Homo sapiens	Human SPROUTY-1 protein, SEQ ID NO:24.	1737	99
49	X04145	Homo sapiens	T3 gamma precursor (aa -22 to 160)	942	99
51	X63547	Homo sapiens	oncogene	5845	99
52	M94043	Rattus norvegicus	rab-related GTP-binding protein	1089	96
53	L31783	Mus musculus	uridine kinase	917	71
54	X83973	Homo sapiens	transcription factor	4486	98
55	AF224741	Homo sapiens	chloride channel protein 7	4128	99
56	W74805	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24.	1491	100
57	Z50907	Homo sapiens	Human TBC-1 cDNA from second transcript.	4824	100
58	D79994	Homo sapiens	similar to ankyrin of Chromatium vinosum.	6089	99
59	D79994	Homo sapiens	similar to ankyrin of Chromatium vinosum.	4014	91
60	Y59738	Homo sapiens	Human normal ovarian tissue derived protein 15.	601	100
61	AB031069	Homo sapiens	protein containing CXXC domain 1	1390	100
62	Y66660	Homo sapiens	Membrane-bound protein PRO783.	2492	99
63	Y66660	Homo sapiens	Membrane-bound protein PRO783.	1709	99
64	S70011	Rattus sp.	tricarboxylate carrier	895	55
65	AF139518	Rattus norvegicus	A-kinase anchor protein	178	24
66	W29666	Homo sapiens	Homo sapiens DH1308_1 clone secreted protein.	157	30
67	AJ245738	Homo sapiens	claudin-15	1206	100
68	AF099138	Rattus norvegicus	GLUT4 vesicle protein	4183	87
69	AF099138	Rattus norvegicus	GLUT4 vesicle protein	4906	86
70	Z82059	Caenorhabditis elegans	Similarity to Drosophila ring canal protein comes from this gene	1285	44
71	AF224278	Homo sapiens	PMEPA1 protein	1282	100
72	AF126426	Homo sapiens	neurotrimin	1809	100
73	Y41652	Homo sapiens	Human MEK2 protein sequence.	2065	99
74	Y41652	Homo sapiens	Human MEK2 protein sequence.	1207	100
75	AF188622	Mus musculus	selectively expressed in embryonic epithelia protein-1	1485	74
76	AE000406	Escherichia coli	putative DNA topoisomerase	950	100
77	X99302	Homo sapiens	Pop1	655	100
78	AL136538	Schizosaccharomyces pombe	similarity to S. cerevisiae kti12 protein	210	31
79	AF129756	Homo sapiens	G4	1554	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
80	AL096768	Homo sapiens	dJ858B16.2 (phosphatidylserine decarboxylase (PSSC, EC 4.1.1.65))	2033	100
81	AL096768	Homo sapiens	dJ858B16.2 (phosphatidylserine decarboxylase (PSSC, EC 4.1.1.65))	1220	96
82	X57351	Homo sapiens	1-8D	677	98
83	AC005594	Homo sapiens	R26984_1	2700	98
84	X73113	Homo sapiens	fast My8P-C	5959	99
85	AF097330	Homo sapiens	H1 chloride channel; p64H1; CLIC4	1305	99
86	AB018423	Mus musculus	SH2 domain-containing protein	1360	78
87	AF272151	Homo sapiens	adaptor protein CIKS	3084	99
88	AF196329	Homo sapiens	triggering receptor expressed on monocytes 1	1214	100
89	AB016879	Arabidopsis thaliana	contains similarity to pre-mRNA splicing factor-gene_id:MRB17.2	634	36
90	AJ133721	Mus musculus	homeodomain protein	654	57
91	AJ242864	Mus musculus	phtf protein	619	61
92	A61971	unidentified	MCSP	11676	99
93	Y99365	Homo sapiens	Human PRO1250 (UNQ633) amino acid sequence SEQ ID NO:86.	3890	100
94	Y87231	Homo sapiens	Human signal peptide containing protein HSPP-8 SEQ ID NO:8.	1031	100
95	AF227741	Rattus norvegicus	protein kinase WNK1	2428	95
96	AF227741	Rattus norvegicus	protein kinase WNK1	1961	94
97	Y92513	Homo sapiens	Human OXRE-10.	1626	100
98	AL021366	Homo sapiens	cICK0721Q.3 (Kinesin related protein)	3423	100
99	AC005783	Homo sapiens	R33083_1	1974	99
100	Y95293	Homo sapiens	Human GEF containing NEK-like kinase substrate 6GNK.	4092	99
101	AL118501	Homo sapiens	dJ1191N16.1 (A novel protein (translation of the cDNA DKFZp566A0946, Em:AL050069))	1509	100
102	AJ006267	Homo sapiens	ClpX-like protein	3233	100
103	AF100753	Homo sapiens	ancient ubiquitous 46 kDa protein AUP1	2042	96
104	AB015982	Homo sapiens	serine/threonine kinase	4718	100
105	AF151074	Homo sapiens	HSPC240	831	64
106	M35522	Canis familiaris	GTP-binding protein (rab7)	354	50
107	R99800	Homo sapiens	NTII-1 nerve protein, facilitates regeneration of nerve cells.	2337	93
108	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1290	93
109	AC005614	Homo sapiens	F23269_2	3369	99
110	AP064729	Homo sapiens	RAN binding protein 16	3285	100
111	X52425	Homo sapiens	interleukin 4 receptor	4496	100
112	Y41686	Homo sapiens	Human PRO274 protein sequence.	2285	100
113	W15506	Homo sapiens	Mitogen activating protein kinase ERK1.	1991	100
114	Y71071	Homo sapiens	Human membrane transport protein, MTRP-16.	1190	99
115	AL049548	Homo sapiens	dJ398G3.1 (ortholog of rat CPG2)	3497	99
116	AF189817	Mus musculus	evectin-2	1124	90
117	W30891	Homo	Human cytostatin III protein.	715	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		sapiens			
118	AF116618	Homo sapiens	PRO1038	1469	100
119	Y08915	Homo sapiens	alpha 4 protein	1748	100
120	AF098070	Drosophila melanogaster	Lis1 homolog	192	39
121	AF052432	Homo sapiens	katanin p80 subunit	181	37
122	Y70743	Homo sapiens	PSEQ-1 protein encoded by NSEQ gene associated with matrix remodelling.	2637	98
123	AF083246	Homo sapiens	HSPC028	2132	100
124	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	833	99
125	M63109	Leishmania major	glycoprotein 96-92	172	27
126	U75467	Drosophila melanogaster	Atu	935	36
127	Z68220	Caenorhabditis elegans	Similarity to Human ADP/ATP carrier protein	438	43
128	AF095927	Rattus norvegicus	protein phosphatase 2C	1927	94
129	W92958	Homo sapiens	Human zsig44 protein.	463	100
130	AF115391	Lactobacillus sakei	ribokinase RbsK	508	37
131	X93498	Homo sapiens	21-Glutamic Acid-Rich Protein	1250	100
132	X93498	Homo sapiens	21-Glutamic Acid-Rich Protein	916	87
133	W52811	Homo sapiens	Human DBI/ACBP -like protein (DBIH).	705	97
134	Y84444	Homo sapiens	Amino acid sequence of a human RNA-associated protein.	3230	100
135	M69181	Homo sapiens	non-muscle myosin B	189	20
136	W74882	Homo sapiens	Human secreted protein encoded by gene 154 clone HE6FL83.	480	100
137	W78200	Homo sapiens	Human secreted protein encoded by gene 75 clone HHGAU81.	855	99
138	AL033520	Homo sapiens	dJ349A12.1 (similar to KIAA0701 protein)	424	39
139	AF020261	Santalum album	proline rich protein	119	30
140	X70394	Homo sapiens	zinc finger protein	1634	100
141	Y06439	Homo sapiens	Human protease HUPM-8.	936	100
142	Z68493	Caenorhabditis elegans	predicted using Genefinder	365	42
143	AB018107	Arabidopsis thaliana	ADP-ribosylation factor-like protein	596	65
144	AF161483	Homo sapiens	HSPC134	580	51
145	Y84902	Homo sapiens	A.human proliferation and apoptosis related protein.	480	100
146	AB004906	Ipomoea purpurea	transposase	146	20
147	AC007357	Arabidopsis thaliana	F3F19.18	647	31
148	W75155	Homo sapiens	Human secreted protein encoded by gene 41 clone HNTME13.	1494	98
149	AF056490	Homo sapiens	cAMP-specific phosphodiesterase 8A	3710	99
150	Y58171	Homo sapiens	Human hydrolase homologue HHH-7.	785	99
151	U10397	Saccharomyces cerevisiae	Yhr148wp	515	53
152	X73478	Homo sapiens	phosphotyrosyl phosphatase activator	1719	99
153	AL049697	Homo sapiens	dJ382I10.5.1 (novel protein	2034	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			similar to arginyl-tRNA)		
154	AF169802	Homo sapiens	cytochrome b5 reductase b5R.2	1455	99
155	X94703	Homo sapiens	rab28	1126	99
156	Y25716	Homo sapiens	Human secreted protein encoded from gene 6.	1471	100
158	W77404	Homo sapiens	Secreted salivary polypeptide zsig32.	937	100
159	Y17248	Homo sapiens	Human protein kinase inhibitor-2 (PKI-2).	383	100
160	J04970	Homo sapiens	carboxypeptidase M precursor	2395	100
161	W54040	Homo sapiens	Human interferon-inducible protein, HIPI.	484	98
162	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)	1357	100
163	AF125535	Homo sapiens	pp21 homolog	193	45
164	G03632	Homo sapiens	Human secreted protein, SEQ ID NO: 7713.	463	97
165	AJ250839	Homo sapiens	serine/threonine protein kinase	1442	71
166	L09649	Zymomonas mobilis	zm2	173	37
167	Y73337	Homo sapiens	HTRM clone 1944530 protein sequence.	1204	100
168	W88645	Homo sapiens	Secreted protein encoded by gene 112 clone HUKFC71.	1084	100
169	AF214731	Homo sapiens	ATP-dependent RNA helicase	4402	100
170	AE000871	Methanobacterium thermoautotrophicum	conserved protein	166	27
171	Y27684	Homo sapiens	Human secreted protein encoded by gene No. 118.	821	100
172	AF226044	Homo sapiens	HSNFRK	2904	100
173	AJ245946	Homo sapiens	neuroglobin	779	100
174	D43949	Homo sapiens	This gene is novel.	3202	100
175	Y07923	Homo sapiens	GTP-binding protein	1205	100
176	W90338	Homo sapiens	Human DP1 homologue protein.	966	100
177	Y41675	Homo sapiens	Human channel-related molecule HCRM-3.	1122	100
178	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.	936	99
179	AF220492	Homo sapiens	krueppel-like zinc finger protein HZF2	4100	99
180	X03084	Homo sapiens	C1q B-chain precursor	1240	100
181	U57344	Mus musculus	Meis3	1813	89
183	U57344	Mus musculus	Meis3	1743	86
184	U57344	Mus musculus	Meis3	1070	86
185	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	1389	58
186	AF200357	Mus musculus	pantothenate kinase 1 beta	1605	82
187	W75058	Homo sapiens	Human secreted protein encoded by gene 2 clone HLDBG33.	1188	99
188	AJ292529	Homo sapiens	suppressor of sterile four 1	2424	100
190	X54134	Homo sapiens	protein-tyrosine phosphatase	3705	100
191	Y22203	Homo sapiens	Human calcium-binding phosphoprotein, CBPP-1, protein sequence.	1083	99
192	W63692	Homo sapiens	Human secreted protein 12.	1975	100
193	W87772	Homo sapiens	Human serum glucocorticoid-regulated kinase (H-SGK2) polypeptide.	2605	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
194	AF084259	Mus musculus	bromodomain-containing protein BP75	693	54
195	Y00752	Rattus norvegicus	serine dehydratase (AA 1 - 327)	994	61
196	W95349	Homo sapiens	Human foetal brain secreted protein fh170_7.	2596	100
197	AB028859	Homo sapiens	hDj9	1890	100
198	W95633	Homo sapiens	Homo sapiens secreted protein gene clone hm236_1.	1614	100
199	Y44277	Homo sapiens	Human nucleic acid methylase-2.	2096	99
200	AB030039	Homo sapiens	hPACPL1	2258	100
201	X54162	Homo sapiens	64 Kd autoantigen	2918	99
202	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
203	X13885	Nicotiana tabacum	extensin (AA 1-620)	185	33
204	J04204	Bos taurus	32 kd accessory protein	1837	100
205	J04204	Bos taurus	32 kd accessory protein	1101	100
207	Y87283	Homo sapiens	Human signal peptide containing protein HSPP-60 SEQ ID NO:60.	1318	100
208	Y02860	Homo sapiens	Fragment of human secreted protein encoded by gene 65.	936	98
209	AL121889	Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))	694	54
210	AF226732	Homo sapiens	NPD007	1345	76
211	X66295	Mus musculus	C1q C chain	970	73
212	Z29328	Homo sapiens	Ubiquitin-conjugating enzyme UbcH2	966	100
213	Z29328	Homo sapiens	Ubiquitin-conjugating enzyme UbcH2	542	98
214	AJ002030	Homo sapiens	progresterone binding protein	1163	100
215	X70649	Homo sapiens	member of DEAD box protein family	3933	100
216	AF250558	Homo sapiens	claudin-2	1169	99
217	AL021453	Homo sapiens	dJ821D11.1 (PUTATIVE protein)	259	100
218	Y08565	Homo sapiens	UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase	3331	99
219	Y94452	Homo sapiens	Human inflammation associated protein	2067	100
220	AL035521	Arabidopsis thaliana	putative protein	315	42
221	AL031786	Schizosaccharomyces pombe	putative proline-trna synthetase	811	41
222	AL109736	Schizosaccharomyces pombe	WD repeat protein	626	40
223	X52493	Glycine max	DNA-directed RNA polymerase	136	23
224	AL035659	Homo sapiens	dJ979N1.1 (dJ979N1.1)	5199	98
225	AB032401	Mus musculus	mmDj4	1761	92
226	AB032401	Mus musculus	mmDj4	1988	92
227	X83502	Saccharomyces cerevisiae	J1007	112	26
228	X83502	Saccharomyces cerevisiae	J1007	79	25
229	AF143723	Homo sapiens	heat shock protein HSP60	2557	99
230	Y66677	Homo sapiens	Membrane-bound protein PRO828.	982	100
231	AB027466	Homo sapiens	spondin 2	1756	99
232	W95634	Homo sapiens	Homo sapiens secreted protein.	1391	100
233	W00365	Homo sapiens	Human cyclin B1.	2218	99
234	Y53762	Homo sapiens	A GTP-binding polypeptide	1017	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			designated RAQ.		
235	Z50749	Homo sapiens	yeast sds22 homolog	1800	100
236	Z50749	Homo sapiens	yeast sds22 homolog	1754	98
237	AB026491	Homo sapiens	PICK1	2137	100
238	AJ270205	Entodinium caudatum	putative phosphatidylinositol-4-phosphate 5-kinase	114	37
239	AB030189	Mus musculus	contains transmembrane (TM) region and ATP binding region	710	93
240	W56538	Homo sapiens	Human hedgehog interacting protein (HIP).	3785	99
241	W56538	Homo sapiens	Human hedgehog interacting protein (HIP).	3436	99
242	AF155107	Homo sapiens	NY-REN-37 antigen	996	99
243	AF155107	Homo sapiens	NY-REN-37 antigen	1005	100
244	AL031320	Homo sapiens	dJ20N2.1 (novel protein similar to yeast and bacterial cytosine deaminase)	763	99
245	U37026	Rattus norvegicus	sodium channel beta 2 subunit	162	30
246	AL078599	Homo sapiens	dJ991C6.1 (novel protein similar to C. elegans F55A12.9 (Tr:P91086))	2391	98
247	U32274	Saccharomyces cerevisiae	Ydr386wp; CAI: 0.12	191	37
248	Y41719	Homo sapiens	Human PRO864 protein sequence.	1879	100
249	AB029434	Homo sapiens	ghrelin precursor	611	100
250	X97831	Rattus norvegicus	carnitine/acylcarnitine carrier protein	246	38
251	W80993	Homo sapiens	Human RIP-interacting factor RIF.	1724	100
252	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
253	W59878	Homo sapiens	Amino acid sequence of the cDNA clone AIF-2 (HEBGM49).	765	100
254	AL354533	Leishmania major	possible adenylate kinase	265	34
255	AF233322	Mus musculus	zinc transporter like 2	1916	95
256	Y78113	Homo sapiens	Human cytokine signal regulator CKSR-1 SEQ ID NO:1.	2247	99
257	AL035539	Arabidopsis thaliana	putative amino acid transport protein	390	27
258	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
259	AL035689	Homo sapiens	dJ187J11.1 (novel protein similar to protein kinase C inhibitors)	974	100
260	AE000909	Methanobacterium thermoautotrophicum	serine/threonine protein kinase related protein	363	30
261	AL050131	Homo sapiens	hypothetical protein	626	100
262	AF019661	Mus musculus	zeta proteasome chain; PSMA5	1214	100
263	AL035593	Homo sapiens	dJ310J6.1 (novel protein)	821	100
264	AL022318	Homo sapiens	bK150C2.3 (PUTATIVE novel protein similar to APOBEC1)	1072	100
265	AF205940	Homo sapiens	endomucin	1289	100
266	AL023583	Homo sapiens	dJ500L14.1 (novel protein)	789	100
267	AL034548	Homo sapiens	dJ1103G7.3 (novel protein kinase domains containing protein similar to phosphoprotein C8FW)	1888	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
268	AF161470	Homo sapiens	HSPC121	1884	98
269	AF161470	Homo sapiens	HSPC121	1232	96
270	X90763	Homo sapiens	HHa5 hair keratin type I intermediate filament	2190	99
271	AF207600	Homo sapiens	ethanolamine kinase	1952	100
272	M32334	Homo sapiens	intercellular adhesion molecule 2	1436	100
273	AF161483	Homo sapiens	HSPC134	663	61
274	Y53052	Homo sapiens	Human secreted protein clone df202_3 protein sequence SEQ ID NO:110.	587	100
276	Y77576	Homo sapiens	Human cytoskeletal protein (HCYT) (clone 2195418).	762	100
277	AF077042	Homo sapiens	30S ribosomal protein S7 homolog	1269	100
278	Y94907	Homo sapiens	Human secreted protein clone ca106_19x protein sequence SEQ ID NO:20.	1619	98
279	Y68788	Homo sapiens	Amino acid sequence of a human phosphorylation effector PHSP-20.	2801	99
280	Z75134	Canis familiaris	rod transducin	1016	100
281	Z75134	Canis familiaris	rod transducin	1718	96
282	AF249873	Homo sapiens	muscle-specific protein	1395	100
283	AL050007	Homo sapiens	hypothetical protein	405	98
284	AF201931	Homo sapiens	DC1	1859	99
285	AF156102	Homo sapiens	ELL complex EAP30 subunit	1318	99
286	Y35897	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 146.	1250	99
287	U88964	Homo sapiens	HEM45	923	100
288	AL050143	Homo sapiens	hypothetical protein	598	100
289	AJ011098	Homo sapiens	telethonin	574	100
290	Y66724	Homo sapiens	Membrane-bound protein PRO836.	2321	100
291	AF034801	Homo sapiens	liprin-alpha4	2565	98
292	AF034801	Homo sapiens	liprin-alpha4	2590	100
293	AL049851	Homo sapiens	dJ889J22B.1 (novel protein (isoform 1))	1738	100
294	Y73348	Homo sapiens	HTRM clone 839651 protein sequence.	1245	99
295	L11672	Homo sapiens	zinc finger protein	1694	44
296	AL035423	Homo sapiens	dJ20I3.1 (brain mitochondrial carrier protein-1 (BMCP1))	1024	79
297	AF198532	Homo sapiens	lymphoid enhancer binding factor-1	2173	100
298	AF161417	Homo sapiens	HSPC299	1147	85
299	AF159141	Homo sapiens	breast cancer metastasis-suppressor 1	1236	99
300	U26397	Rattus norvegicus	inositol polyphosphate 4-phosphatase	160	30
301	AF036145	Homo sapiens	meningioma-expressed antigen 5	3458	100
302	Z82022	Homo sapiens	GlcNac-1-P transferase	2067	99
303	AF269232	Mus musculus	butyrophilin-like protein BUTR-1	271	50
304	AJ222644	Arabidopsis thaliana	asparaginyl-tRNA synthetase	659	50
305	AF054180	Homo sapiens	hematopoietic cell derived zinc finger protein	351	79
306	AJ272079	Homo sapiens	APOBEC-1 stimulating protein	3056	100
308	Y44486	Homo sapiens	Human GPRW receptor polypeptide.	1721	100
309	AJ131891	Homo sapiens	DNA polymerase mu	2598	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
310	AF293335	Homo sapiens	p30 DBC	1248	92
311	AF176525	Mus musculus	F-box protein FBL12	1501	93
312	X57802	Homo sapiens	Immunoglobulin lambda light chain	959	81
313	Z36715	Homo sapiens	Net	2048	98
314	AF161532	Homo sapiens	HSPC047	727	100
315	AF208068	Homo sapiens	kelch-like protein KLHL3a	3046	100
316	Y66666	Homo sapiens	Membrane-bound protein PRO1013.	1166	100
317	Y29666	Homo sapiens	Human Ras protein RAPR-1.	1253	98
318	AJ387747	Homo sapiens	sialin	2614	99
319	AF161362	Homo sapiens	HSPC099	224	40
320	Y68773	Homo sapiens	Amino acid sequence of a human phosphorylation effector PHSP-5.	2243	99
321	AJ238379	Homo sapiens	putative TH1 protein	3013	100
322	AB040812	Homo sapiens	protein kinase PAK5	3792	99
323	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	913	100
324	Y13381	Homo sapiens	Amino acid sequence of protein PRO271.	1976	100
325	Y94944	Homo sapiens	Human secreted protein clone bf157_16 protein sequence SEQ ID NO:94.	2305	98
326	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	6728	99
327	AF198532	Homo sapiens	lymphoid enhancer binding factor-1	2173	100
328	Z78013	Caenorhabditis elegans	Similarity to Drosophila Cadherin-related tumor suppressor	569	33
329	AF212921	Mus musculus	MMTV receptor variant 1	484	94
330	Z75330	Homo sapiens] >R65207 R65207 02-MAR-1995 27-AUG-1993 Human stromalin-1. [Homo sapiens	nuclear protein SA-1	6492	99
331	AL008583	Homo sapiens	dJ327J16.3 (supported by GENSCAN, PGENES and GENENISE)	2133	99
332	Y36104	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 489.	310	41
333	AJ271669	Homo sapiens	putative sialoglycoprotease	1747	100
334	AF156598	Mus musculus	p53-regulated DDA3	997	64
335	M99058	Eimeria maxima	em100 gene is homologous the Eimeria tenella gene et100	154	26
336	Y85564	Homo sapiens	Human homologue of UNC-53 (Hs-UNC-53/1) sequence.	3386	97
337	Y85564	Homo sapiens	Human homologue of UNC-53 (Hs-UNC-53/1) sequence.	2602	94
338	Y85564	Homo sapiens	Human homologue of UNC-53 (Hs-UNC-53/1) sequence.	3447	98
339	Z66561	Caenorhabditis elegans	Similarity to Human rab13 protein (PIR Acc. No. A49647).	716	34
340	AB021643	Homo sapiens	gonadotropin inducible transcription repressor-3	2761	99
341	G01946	Homo sapiens	Human secreted protein, SEQ ID NO: 6027.	465	98
342	AF020591	Homo sapiens	zinc finger protein	1091	48
343	L29154	Homo sapiens	Immunoglobulin heavy chain	439	84

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			VDJ region		
344	U10281	Sus scrofa	gastric mucin	279	24
345	AK000404	Homo sapiens	unnamed protein product	1177	99
346	L22557	Rattus norvegicus	calmodulin-binding protein	1949	84
347	L22557	Rattus norvegicus	calmodulin-binding protein	2363	91
348	AL049481	Arabidopsis thaliana	AIG1-like protein	316	30
350	AJ251516	Mus musculus	cysteine and histidine-rich protein	1460	99
351	AK024477	Homo sapiens	FLJ00070 protein	1773	100
352	U50133	Homo sapiens	ankyrin	502	33
353	AK000625	Homo sapiens	unnamed protein product	721	100
354	AF161420	Homo sapiens	HSPC302	2623	97
355	AJ010014	Homo sapiens	M96A protein	1269	47
356	AF151029	Homo sapiens	HSPC195	941	91
357	AL022327	Homo sapiens	dJ355C18.1 (KIAA0027)	1911	100
358	W78128	Homo sapiens	Human secreted protein encoded by gene 3 clone HOSBI96.	1117	100
359	X03414	Drosophila melanogaster	Kr polypeptide	316	45
360	AF151079	Homo sapiens	HSPC245	643	100
361	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	530	41
362	AF254741	Drosophila melanogaster	Centaurin Gamma 1A	681	46
363	AF213465	Homo sapiens	dual oxidase	2016	100
364	AF181562	Homo sapiens	proSAAS	1319	100
365	AF181562	Homo sapiens	proSAAS	1024	99
366	U73200	Mus musculus	p116Rip	884	82
367	AF263744	Homo sapiens	erbb2-interacting protein ERBIN	4973	99
368	U37501	Mus musculus	laminin alpha 5 chain	5867	72
369	AF043695	Caenorhabditis elegans	similar to the protein phosphates 2c family	549	36
370	Y73440	Homo sapiens	Human secreted protein clone yj23_1 protein sequence SEQ ID NO:102.	1484	99
371	AF272833	Homo sapiens	misato	2869	97
372	AF198454	Homo sapiens	epithelial protein lost in neoplasm beta	3927	100
373	Y73345	Homo sapiens	HTRM clone 438283 protein sequence.	273	80
374	AF169017	Homo sapiens	formiminotransferase cyclodeaminase	2717	98
375	A95106	unidentified	RED ALPHA	1202	99
376	W74828	Homo sapiens	Human secreted protein encoded by gene 100 clone HLQA952.	1012	99
377	Y32131	Homo sapiens	Human LYST-2 protein.	3556	99
378	M14912	Homo sapiens	pol	132	86
379	AF090934	Homo sapiens	PRO0518	382	100
380	X66363	Homo sapiens	serine/threonine protein kinase	2499	100
381	Y41699	Homo sapiens	Human PRO703 protein sequence.	2362	100
382	AF174498	Homo sapiens	GR AF-1 specific protein phosphatase	7008	98
383	U64608	Caenorhabditis elegans	coded for by C. elegans cDNA yk173c12.5	246	36
384	U50133	Homo sapiens	ankyrin	502	33
385	AJ238520	Homo sapiens	putative transcription factor-like nuclear regulator	4123	97

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
387	AF208845	Homo sapiens	BM-003	1375	99
389	X57821	Homo sapiens	immunoglobulin lambda light chain	797	76
390	AF182404	Homo sapiens	mitochondrial uncoupling protein 1	1670	99
391	Y85564	Homo sapiens	Human homologue of UNC-53 (Hs-UNC-53/1) sequence.	3386	97
393	AF178432	Homo sapiens	SH3 protein	3700	100
394	AF229928	Drosophila melanogaster	cytoplasmic protein 89BC	1616	62
395	AF181721	Homo sapiens	RU2S	2254	100
396	Y69197	Homo sapiens	Amino acid sequence of a human betaIV-spectrin protein.	1626	98
397	U48238	Mus musculus	zinc finger protein neuro-d4	749	60
398	AL390137	Homo sapiens	hypothetical protein	263	51
399	AF217525	Homo sapiens	Down syndrome cell adhesion molecule	5337	60
400	AL022599	Schizosaccharomyces pombe	WD repeat protein	447	27
401	AC004859	Homo sapiens	similar to 2-oxoglutarate dehydrogenase ; similar to Q02218 (PID:g1352618)	4176	78
402	AB010266	Mus musculus	tenascin-X	10246	62
403	AL133288	Homo sapiens	dJ671D7.1 (similar to D.melanogaster CG5986 protein)	761	100
404	Z68753	Caenorhabditis elegans	ZC518.3b	888	48
405	Z78013	Caenorhabditis elegans	Similarity to Drosophila Cadherin-related tumor suppressor	569	33
406	AB031230	Homo sapiens	protein containing CXXC domain 2	1196	97
407	AF155106	Homo sapiens	NY-REN-36 antigen	1168	100
408	Y57945	Homo sapiens	Human transmembrane protein HTMPN-69.	1538	99
409	Z18361	Ovis aries	trichohyalin	184	30
410	AF249744	Homo sapiens	RhoGEF	2733	100
411	AF176529	Mus musculus	F-box protein FBX13	2072	94
412	AF210842	Homo sapiens	HARP	4880	100
413	AL031658	Homo sapiens	dJ310O13.7 (novel protein similar to H. roretzi HRPET-3)	776	98
414	X57398	Homo sapiens	pm5 protein	6131	99
415	AB029826	Homo sapiens	3-methylcrotonyl-CoA carboxylase biotin-containing subunit	2961	99
416	U43503	Saccharomyces cerevisiae	Lph1p	115	42
417	AL160493	Leishmania major	possible t26f17.21	239	35
418	Y08100	Homo sapiens	Human PRO331 protein.	330	29
419	U15131	Homo sapiens	p126	2228	54
420	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	2363	100
421	AF190635	Drosophila melanogaster	ankyrin 2	755	30
422	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	1962	100
423	AL137530	Homo sapiens	hypothetical protein	433	94
424	X63753	Homo sapiens	son-a	7269	100
425	AB027249	Homo sapiens	MAPKK like protein kinase	1693	100
426	AF279144	Homo sapiens	tumor endothelial marker 7 precursor	1084	55

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
427	AF279144	Homo sapiens	tumor endothelial marker 7 precursor	1259	56
428	AE003683	Drosophila melanogaster	CG8312 gene product	149	29
429	Y07829	Homo sapiens	RING finger protein	2201	99
430	AF096897	Drosophila melanogaster	pushover	4442	47
431	U41387	Homo sapiens	Gu protein	4021	99
432	AF023674	Homo sapiens	nephrocystin	3783	100
433	AF146760	Homo sapiens	septin 2-like cell division control protein	2284	100
434	AB006697	Arabidopsis thaliana	cleft lip and palate associated transmembrane protein-like	886	42
437	Y94247	Homo sapiens	Human calcium binding protein hCBP.	1704	100
438	AB040672	Homo sapiens	UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase	1075	63
439	AF105228	Bos taurus	tuftelin	285	33
440	R06463	Homo sapiens	Derived protein of clone ICA13 (ATCC 40553).	3073	99
441	X14971	Mus musculus	alpha-adaptin (A) (AA 1-977)	4897	98
442	X53773	Rattus norvegicus	alpha-c large chain (AA 1-938)	3979	81
443	Y66689	Homo sapiens	Membrane-bound protein PRO1136.	3299	99
444	AC067754	Arabidopsis thaliana	unknown protein; 20348-23707	114	33
445	AF229032	Mus musculus	pIL	2077	93
446	AF056035	Rattus norvegicus	s-nexilin	2662	85
447	AF132484	Mus musculus	unknown	478	51
448	W89024	Homo sapiens	Polypeptide fragment encoded by gene 156.	528	45
449	AF161445	Homo sapiens	HSPC327	1606	100
450	Z68753	Caenorhabditis elegans	ZC518.3b	951	49
451	W39160	Homo sapiens	Human partial complement factor H protein fragment 3.	155	32
452	W85727	Homo sapiens	Novel protein (Clone BM46_10).	2799	99
453	Y53629	Homo sapiens	A bone marrow secreted protein designated BMS115.	2810	100
454	D87438	Homo sapiens	Similar to a C.elegans protein in cosmid C14H10	4069	100
455	AF240468	Homo sapiens	nicastrin	3687	100
456	Z15005	Homo sapiens	CENP-E	13305	99
457	M59216	Homo sapiens	gamma-aminobutyric acid receptor beta-1 subunit	2477	100
458	Y73467	Homo sapiens	Human secreted protein clone yd61_1 protein sequence SEQ ID NO:156.	966	100
459	W67824	Homo sapiens	Human secreted protein encoded by gene 18 clone HSLFM29.	535	100
460	AF163151	Homo sapiens	dentin sialophosphoprotein precursor	279	19
461	D87446	Homo sapiens	Similar to a C.elegans protein encoded in cosmid C27F2 (U40419)	9196	99
462	G04044	Homo sapiens	Human secreted protein, SEQ ID NO: 8125.	486	93
463	AC002398	Homo sapiens	F25965_1	1018	100
464	AF064856	Rattus sp.	7acomp protein	1845	84
465	AF223408	Homo sapiens	B99	3686	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
466	AF223408	Homo sapiens	B99	2878	87
467	AF104415	Mus musculus	gene trap locus-13	6336	91
468	U53450	Rattus norvegicus	Jun dimerization protein 1 JDP-1	196	49
469	AL031297	Homo sapiens	dJ97P20.1 (novel gene)	3564	99
470	AF257077	Homo sapiens	eukaryotic translation initiation factor EIF2B subunit 3	1274	95
471	L28125	Podospora anserina	beta transducin-like protein	284	38
472	Y84903	Homo sapiens	A human proliferation and apoptosis related protein.	2337	100
473	AF144237	Homo sapiens	LOMP protein	252	44
474	Y71213	Homo sapiens	Human irritable bowel disease related polypeptide IMX39.	838	100
475	Y95006	Homo sapiens	Human secreted protein ve13_1, SEQ ID NO:52.	3411	100
476	D38549	Homo sapiens	ha1025 is new	6533	99
477	AF241230	Homo sapiens	TAK1-binding protein 2	3656	100
478	AL031534	Schizosaccharomyces pombe	putative asparagine synthase	482	40
479	L28125	Podospora anserina	beta transducin-like protein	233	26
480	AF161544	Homo sapiens	HSPC059	434	77
481	AJ238248	Homo sapiens	centaurin beta2	3986	99
482	Z38061	Saccharomyces cerevisiae	mal5, sta1, len: 1367, CAI: 0.3, AMYH_YEAST P08640 GLUCOAMYLASE S1 (EC 3.2.1.3)	295	23
483	AF161381	Homo sapiens	HSPC263	1404	100
484	AF223468	Homo sapiens	AD021 protein	1314	100
486	X57527	Homo sapiens	alpha 1(VIII) collagen	4166	99
487	Y19062	Homo sapiens	39k3 protein	2475	100
488	Y73373	Homo sapiens	HTRM clone 921803 protein sequence.	555	56
489	AL021918	Homo sapiens	b3418.1 (Kruppel related zinc Finger protein 184)	4184	100
490	X53773	Rattus norvegicus	alpha-c large chain (AA 1-938)	4675	97
491	U52426	Homo sapiens	GOK	1459	59
492	AL359773	Leishmania major	possible threonine synthase	702	45
493	AF226614	Homo sapiens	ferroportin1	2929	100
494	Z93241	Homo sapiens	dJ222E13.1 (novel protein with some similarity to Drosophila KRAKEN)	513	96
495	AF036977	Homo sapiens	unknown	1812	100
496	U93564	Homo sapiens	p40	133	45
497	Y91405	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:126.	357	100
498	AF069781	Drosophila melanogaster	Bem46-like protein	653	43
499	Y16601	Homo sapiens	Human cell-cycle phosphoprotein CECYP-2.	1658	98
500	X70944	Homo sapiens	PTB-associated splicing factor	3883	100
501	AF027503	Mus musculus	putative membrane-associated guanylate kinase 1	205	36
502	AF282874	Homo sapiens	nectin 3; PRR3	2856	99
503	AJ249732	Homo sapiens	G8 protein	669	100
504	AF208861	Homo sapiens	BM-019	1629	100
505	L09708	Homo sapiens	complement component C2	4022	100
507	X66285	Mus musculus	HC1 ORF	115	43
508	D00189	Rattus norvegicus	Na+,K+-ATPase alpha-subunit	5227	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
509	Y94971	Homo sapiens	Human secreted protein clone fa171_1 protein sequence SEQ ID NO:148.	2176	100
510	AB019038	Homo sapiens	beta-1,4 mannosyltransferase	781	77
511	AB019038	Homo sapiens	beta-1,4 mannosyltransferase	1347	100
512	AB019038	Homo sapiens	beta-1,4 mannosyltransferase	1520	99
513	X84908	Homo sapiens	phosphorylase kinase	5729	99
514	X52851	Homo sapiens	peptidylprolyl isomerase	650	76
515	AF186084	Homo sapiens	epidermal growth factor repeat containing protein	3046	99
516	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7683.	505	99
517	U04706	Bos taurus	50 kDa protein	1749	77
518	G00653	Homo sapiens	Human secreted protein, SEQ ID NO: 4734.	530	100
519	AF161475	Homo sapiens	HSPC126	1368	100
520	Y99366	Homo sapiens	Human PRO1475 (UNQ746) amino acid sequence SEQ ID NO:88.	3394	97
521	AF266852	Homo sapiens	PTPLA	1295	100
522	AE000995	Archaeoglobus fulgidus	chromosome segregation protein (smc1)	153	20
523	AF062249	Homo sapiens	immunoglobulin heavy chain variable region	605	97
524	AJ223830	Rattus norvegicus	ARE1	2950	98
525	W01535	Homo sapiens	Cellular homologue of the SV40 large T antigen.	1276	83
526	AF145658	Drosophila melanogaster	BcDNA.GH10229	320	33
527	AF112213	Homo sapiens	putative Rab5-interacting protein	524	79
528	D49387	Homo sapiens	NADP dependent leukotriene b4 12-hydroxydehydrogenase	1616	100
529	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	328	32
530	AL079335	Homo sapiens	dJ132F21.3 (72.1 KDa protein (DKFZP564A032, SBBI88) similar to mouse IFN-gamma induce MG11.)	1059	99
531	Y91506	Homo sapiens	Human secreted protein sequence encoded by gene 56 SEQ ID NO:179.	1159	98
532	X76116	Caenorhabditis elegans	carrier protein (c2)	576	50
533	X76116	Caenorhabditis elegans	carrier protein (c2)	506	50
534	X12966	Homo sapiens	3-oxoacyl-CoA thiolase propeptide (424 AA)	1972	100
535	Y09267	Homo sapiens	flavin-containing monooxygenase 2	2486	100
536	Z11773	Homo sapiens	SRE-ZBP	2201	99
537	D84224	Homo sapiens	methionyl tRNA synthetase	4741	99
538	D84224	Homo sapiens	methionyl tRNA synthetase	3887	99
539	D84224	Homo sapiens	methionyl tRNA synthetase	2933	96
540	D84224	Homo sapiens	methionyl tRNA synthetase	4529	99
541	J03244	Bos taurus	H+ ATPase 31kDa subunit (EC 3.6.1.3)	848	77
542	Y92514	Homo sapiens	Human OXRE-11.	2301	99
543	AF221712	Homo sapiens	Smad- and Olf-interacting zinc finger protein	2151	61
544	AE000919	Methanobacterium thermoautotrophicum	conserved protein	207	38
545	A06669	synthetic construct	preTGF-beta1	2070	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
546	Y02698	Homo sapiens	Human secreted protein encoded by gene 49 clone HTPCS60.	854	98
547	AF112205	Homo sapiens	WSB-1 protein	2275	100
548	X60271	Mus musculus	c-rel	2264	74
549	AC016827	Arabidopsis thaliana	putative GTPase	810	42
550	Y70400	Homo sapiens	Human cell-signalling protein-2.	429	68
551	AB040365	Homo sapiens	NEDD4-like ubiquitin ligase 1	8290	99
552	Y57880	Homo sapiens	Human transmembrane protein HTMPN-4.	1112	95
553	AF119855	Homo sapiens	PRO1847	265	67
554	M17236	Homo sapiens	MHC HLA-DQ alpha precursor	1332	100
555	AL078468	Arabidopsis thaliana	putative protein	540	40
556	AC006963	Homo sapiens	similar to Kelch proteins; similar to BAA77027 (PID:g4650844)	515	44
557	AK024487	Homo sapiens	FLJ00086 protein	1623	98
558	M12140	Homo sapiens	pol gene protein; Xxx	117	48
559	W74825	Homo sapiens	Human secreted protein encoded by gene 97 clone HAQBF73.	225	56
560	X56681	Homo sapiens	junD protein	373	88
561	AF003136	Caenorhabditis elegans	contains weak similarity to an AMP-binding motif	2926	54
562	AL109839	Homo sapiens	dJ1069P2.3.1 (novel PABPC1 (poly(A)-binding protein)	877	100
563	AF181640	Drosophila melanogaster	BcDNA.GH09817	289	42
564	AF052723	Feline leukemia virus	gag-pol precursor polyprotein gPr80	1547	43
565	AF161472	Homo sapiens	HSPC123	439	44
566	Y28817	Homo sapiens	pt326_4 secreted protein.	3338	100
567	U09848	Homo sapiens	zinc finger protein	1738	100
569	AF155113	Homo sapiens	NY-REN-55 antigen	3603	93
570	AF155113	Homo sapiens	NY-REN-55 antigen	3951	99
571	AL032821	Homo sapiens	dJ55C23.1 (vanin 1)	1821	98
572	M69181	Homo sapiens	non-muscle myosin B	7350	99
573	M69181	Homo sapiens	non-muscle myosin B	7311	98
574	Y59678	Homo sapiens	Secreted protein 108-008-5-0-B6-FL.	772	100
575	AL365234	Arabidopsis thaliana	putative protein	788	40
576	AL365234	Arabidopsis thaliana	putative protein	788	40
577	X06745	Homo sapiens	DNA polymerase alpha-subunit (AA 1 - 1462)	7619	99
578	AB041642	Homo sapiens	PAR-6	1342	100
579	D86984	Homo sapiens	similar to yeast adenylate cyclase (S56776)	2446	100
580	AF165124	Homo sapiens	gamma-aminobutyric acid A receptor gamma 2	2499	99
581	W88812	Homo sapiens	Polypeptide fragment encoded by gene 58.	2339	99
582	U82319	Homo sapiens	novel ORF	342	100
583	P92219	Homo sapiens (human)	CR1 protein.	11425	99
584	AJ223948	Homo sapiens	RNA helicase	6608	99
585	Y08612	Homo sapiens	88kDa nuclear pore complex protein	3874	99
586	Y42384	Homo sapiens	Amino acid sequence of lv310_7.	1007	37
587	AF129756	Homo sapiens	BAT4	1873	98

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
588	AF131775	Homo sapiens	Unknown	1929	99
589	AJ250865	Homo sapiens	TESS 2	2348	100
591	Z98885	Homo sapiens	dJ522J7.2 (bromodomain-containing 1 (similar to peregrin, BR140))	4167	100
592	L76571	Homo sapiens	nuclear hormone receptor	1355	100
593	AF091622	Homo sapiens	PHD finger protein 3	9054	100
594	X56807	Homo sapiens	desmocollin type 2a	4443	100
595	AL137802	Homo sapiens	dJ798A10.1 (novel protein)	212	55
596	AL022329	Homo sapiens	bK407F11.2 (adrenergic, beta, receptor kinase 2)	3653	100
597	AF226048	Homo sapiens	GL003	2009	99
598	AJ278112	Homo sapiens] >Y49635 Y49635 21-OCT-1999 15-APR-1998 Human sdp3.5 protein. [Homo sapiens	putative cell cycle control protein	335	23
599	Y59741	Homo sapiens	Human normal ovarian tissue derived protein 18.	1574	99
600	L36531	Homo sapiens	integrin alpha 8 subunit	5386	99
601	Y38458	Homo sapiens	Human secreted protein encoded by gene No. 20.	895	100
602	AF218584	Homo sapiens	GGA1	3265	100
603	Y13115	Homo sapiens	serine/threonine protein kinase	5071	99
604	AL132776	Homo sapiens	dJ393D12.1 (KIAA0776)	2413	99
605	AL034452	Homo sapiens	dJ682J15.1 (novel Collagen triple helix repeat containing protein)	1979	100
606	Y14494	Homo sapiens	aralar1	3465	99
607	AJ001981	Homo sapiens	OXA1L	2603	100
608	X86098	Homo sapiens	binds directly to adenovirus type 5 E1A protein	3069	100
610	AF163572	Homo sapiens	Forssman glycolipid synthetase	1865	99
611	AF161503	Homo sapiens	HSPC154	1261	97
612	L41834	Ensis minor	nuclear protein	345	30
613	Y91954	Homo sapiens	Human cytoskeleton associated protein 9 (CYSKP-9).	3668	100
614	AL022327	Homo sapiens	dJ355C18.1 (KIAA0027)	361	94
615	X85786	Homo sapiens	binding regulatory factor	3203	100
616	Y08319	Homo sapiens	kinesin-2	3487	99
617	D12644	Mus musculus	KIF2 protein	3609	97
618	U28789	Mus musculus	PACT	5936	89
619	Y35914	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 163.	1684	99
620	AB046382	Mus musculus	testis-abundant finger protein	199	23
621	Y00062	Homo sapiens	precursor polypeptide (AA -23 to 1120)	3440	99
622	AF068286	Homo sapiens	HDCMD38P	861	100
623	X98248	Homo sapiens	sortilin	4436	99
624	X61100	Homo sapiens	75 kDa subunit NADH dehydrogenase precursor	3734	99
625	S58544	Homo sapiens	75 kda infertility-related sperm protein	2125	99
626	AF151027	Homo sapiens	HSPC193	582	93
627	X14968	Homo sapiens	RII-alpha subunit (AA 1-404)	2079	100
628	Y50911	Homo sapiens	Human fetal brain cDNA clone vb7_1 derived protein	1983	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
629	Y50911	Homo sapiens	Human fetal brain cDNA clone vb7_1 derived protein	1694	100
630	AF098786	Homo sapiens	17 beta-hydroxysteroid dehydrogenase type VII	1754	100
631	AL034555	Homo sapiens	dJ134O19.3 (zinc finger protein 151 (pHZ-67))	4273	100
632	W74826	Homo sapiens	Human secreted protein encoded by gene 98 clone HAQBT94.	794	96
633	AF288288	Homo sapiens	HPT protein	2236	100
634	AF041429	Homo sapiens	pRGR1	823	99
635	X66357	Homo sapiens	serine/threonine protein kinase	1589	100
636	Y11284	Homo sapiens	AFX1	2571	98
637	AB004884	Homo sapiens	PKU-alpha	3718	99
638	AJ002303	Homo sapiens	synaptogyrin 1c	1020	100
639	AJ002304	Homo sapiens	synaptogyrin 1b	1002	100
640	AJ002303	Homo sapiens	synaptogyrin 1c	933	94
641	D87682	Homo sapiens	similar to a C.elegans protein encoded in cosmid T26A5.	2676	100
642	M14660	Homo sapiens	ISG-K54	2473	99
643	X06661	Homo sapiens	calbindin (AA 1-261)	1358	100
644	AF119900	Homo sapiens	PRO2822	185	76
645	AB031048	Drosophila melanogaster	microtubule associated-protein orbit	738	27
646	AF250842	Drosophila melanogaster	multiple asters	834	29
647	X86691	Homo sapiens	Mi-2 protein	10110	99
648	U67934	Homo sapiens	44.9 kDa protein C18B11 homolog	827	96
649	AF236061	Oryctolagus cuniculus	RING-finger binding protein	3830	91
650	AL034553	Homo sapiens	dJ914P20.2 (KIAA0784 protein similar to Mus musculus activity-dependent neuroprotective protein (Admp))	5708	100
653	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2388	99
654	AC004614	Homo sapiens	similar to f-spondin proteins AB006086 (PID:g2529225)	3026	99
655	Y57908	Homo sapiens	Human transmembrane protein HTMPN-32.	608	99
656	Z34975	Homo sapiens	ldlCp	3733	100
658	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	1942	99
659	W76734	Homo sapiens	Human mDia Rho targeting protein.	781	34
660	AF202724	Homo sapiens	Sad1 unc-84 domain protein 1	2172	100
661	Z21966	Homo sapiens	mPOU homeobox protein	1529	100
662	AJ242954	Mus musculus	dysferlin	4752	59
663	AF182316	Homo sapiens	myoferlin	6232	99
665	AL161516	Arabidopsis thaliana	hypothetical protein	209	30
667	X59303	Homo sapiens	valyl-tRNA synthetase	3393	99
668	Y13355	Homo sapiens	Amino acid sequence of protein PRO220.	3692	100
669	AB010692	Arabidopsis thaliana	contains similarity to endo-beta-N-acetylglucosaminidase gene	611	52
671	X56123	Mus musculus	talin	4474	76
672	AB039371	Homo sapiens	mitochondrial ABC transporter 3	2902	99
673	AF269223	Homo sapiens	TCP11	806	42
674	AF229633	Mus musculus	groucho-related protein 4	4053	99
675	L14463	Rattus	'transducin	3619	92

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		norvegicus			
676	AC005757	Homo sapiens	R32611_1	2779	100
677	S61069	Homo sapiens	reverse transcriptase homolog=pol (retroviral element)	252	65
678	AF271388	Homo sapiens	CMP-N-acetylneuraminic acid synthase	2273	100
679	X79066	Homo sapiens	ERF-1	1783	100
680	AF118566	Mus musculus	hematopoietic zinc finger protein	769	50
681	Y51415	Homo sapiens	Human wild type pKe83 protein.	2621	99
682	AL133545	Homo sapiens	bA386N14.1 (novel protein similar to a dual specificity phosphatase)	700	68
683	Y86214	Homo sapiens	Nuclear transport protein clone hfb341 protein sequence.	5888	99
684	Y94952	Homo sapiens	Human secreted protein clone fh116_11 protein sequence SEQ ID NO:110.	354	98
685	AL021878	Homo sapiens	dJ257I20.4 (transcription factor 20 (AR1) (KIAA0292) (isoform 2))	154	67
686	AE000198	Escherichia coli	orf, hypothetical protein	628	100
687	M58378	Homo sapiens	synapsin I	3730	99
688	AF039697	Homo sapiens	antigen NY-CO-31	508	98
689	U09355	Oryctolagus cuniculus	protein phosphatase 2A1 B gamma subunit	2356	99
690	AF155106	Homo sapiens	NY-REN-36 antigen	265	50
691	AC004774	Homo sapiens	Dlx-5	1542	100
692	X90530	Homo sapiens	ragB	1926	99
693	X90530	Homo sapiens	ragB	1405	99
694	X90530	Homo sapiens	ragB	1590	85
695	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
696	AC011810	Arabidopsis thaliana	Putative methionine aminopeptidase	669	52
697	AJ250425	Rattus norvegicus	Collybistin I	2455	98
698	AB037901	Homo sapiens	gene amplified in squamous cell carcinoma-1	5364	99
699	Y99401	Homo sapiens	Human PRO1327 (UNQ687) amino acid sequence SEQ ID NO:218.	1386	100
701	AF221712	Homo sapiens	Smad- and Olf-interacting zinc finger protein	6705	100
702	X83573	Homo sapiens	ARSE	3184	99
703	AJ243274	Homo sapiens	AP-2rep protein	2078	99
704	Y71262	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1697	94
705	Y71262	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1736	99
706	Y41257	Homo sapiens	Amino acid sequence of long human FAIM.	1060	100
707	AL022237	Homo sapiens	bK1191B2.3 (PUTATIVE novel Acyl Transferase similar to C. elegans C50D2.7) (isoform 1))	2030	100
708	AJ006266	Homo sapiens	AND-1 protein	5942	100
709	G01571	Homo sapiens	Human secreted protein, SEQ ID NO: 5652.	777	99
710	Y08698	Homo sapiens	ranbp3	2849	98
711	Y68770	Homo sapiens	Amino acid sequence of a human phosphorylation effector PHSP-2.	754	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
712	U93574	Homo sapiens	putative p150	799	59
713	AC004531	Homo sapiens	Gene with similaity to DEAD box helicases	2715	99
714	D89016	Homo sapiens	Neuroblastoma	538	48
715	Y92175	Homo sapiens	Human cardiovascular system associated protein tyrosine phosphatase 2.	734	98
716	AL137013	Homo sapiens	bA311P8.3 (probable uracil phosphoribosyltranferase)	862	100
717	AB035123	Mus musculus	GD1 alpha/GT1a alpha/GQ1b alpha synthase	1696	93
718	Y96290	Homo >P40254 P40254 25-OCT-1984 09-APR-1983 Human IgD. [Homo sapiens	Human IGFAM-2 immunoglobulin.	2345	85
719	X07979	Homo sapiens	integrin beta 1 subunit precursor	4347	99
720	AJ224819	Homo sapiens	tumor suppressor	2149	99
721	Y07595	Homo sapiens	transcription factor TFIIH	2373	100
722	W41565	Homo sapiens] >W41564 W41564 08-OCT-1997 05-APR-1996 Human calpain. [Homo sapiens	Human calpain.	1591	99
723	AF161341	Homo sapiens	HSPC078	1097	98
724	AF187318	Homo sapiens	F-box protein Fbx2	1607	100
725	AC006708	Caenorhabdit is elegans	contains simlarity to Saccharomyces cerevisiae pre-mRNA splicing protein PRP31 (GB:Z72876)	1143	46
726	AC006708	Caenorhabdit is elegans	contains simlarity to Saccharomyces cerevisiae pre-mRNA splicing protein PRP31 (GB:Z72876)	988	46
727	AC024818	Caenorhabdit is elegans	contains similarity to Pfam family PF00400 (WD domain, G-beta repeat), score=81.8, E=1.4e-20, N=3	950	44
728	AJ005897	Homo sapiens	JM5	831	47
729	Y45377	Homo sapiens	Human secreted protein fragment encoded from gene 27.	908	97
730	G03931	Homo sapiens	Human secreted protein, SEQ ID NO: 8012.	578	100
731	AB012720	Oncorhynchus masou	GTP-binding protein	3865	76
732	W73404	Homo sapiens	Human secreted protein encoded by Gene No. 8.	862	97
733	G02650	Homo sapiens	Human secreted protein, SEQ ID NO: 6731.	644	97
734	AC024813	Caenorhabdit is elegans	Hypothetical protein Y54F10AL.a	152	24
735	AL035461	Homo sapiens	dJ967N21.6 (novel CDP-alcohol phosphatidyltransferase family member protein)	1562	98
736	U00033	Caenorhabdit is elegans	similar to S. cerevisiae YJU2 protein	605	41
737	AF079098	Homo sapiens	arginine-tRNA-protein transferase 1-1p; ATE1-1p	2733	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
738	AJ131712	Homo sapiens	nucleolar RNA-helicase	2793	100
739	AJ133115	Homo sapiens	TSC-22-like protein	2054	99
740	X98258	Homo sapiens	M-phase phosphoprotein 9	953	100
741	X98258	Homo sapiens	M-phase phosphoprotein 9	564	74
742	U97191	Caenorhabditis elegans	strong similarity to the YPT1 sub-family of RAS proteins	960	85
743	X76057	Homo sapiens	phosphomannose isomerase	2191	100
744	G03209	Homo sapiens	Human secreted protein, SEQ ID NO: 7290.	496	98
745	X97064	Homo sapiens	Sec23 protein	4034	99
746	W93946	Homo sapiens	Human regulatory molecule HRM-2 protein.	994	100
747	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	1565	99
748	M19529	Sus scrofa	follistatin A	1906	98
749	AJ249457	Trichomonas vaginalis	centrin, putative	183	28
750	AC004410	Homo sapiens	fos39554_1	2094	100
751	AF074968	Homo sapiens	p47ING3 protein	2167	100
752	AF252284	Homo sapiens	transcription specificity factor Sp1	4005	100
753	AB049629	Homo sapiens	phospholysine phosphohistidine inorganic pyrophosphate phosphatase	1375	99
754	D79205	Homo sapiens	ribosomal protein L39	160	77
755	AB008430	Homo sapiens	CDEP	142	29
758	L32162	Homo sapiens	transcription factor	574	80
759	AF037204	Homo sapiens	RING zinc finger protein	295	54
760	Y44250	Homo sapiens	Human cell signalling protein-13.	625	100
761	AF218586	Homo sapiens	Cide-b	1136	100
762	U38934	Gallus gallus	histone H2A	625	97
763	AF226053	Homo sapiens	HSKM-B	606	32
764	X13403	Homo sapiens	Oct-1 protein (AA 1 - 743)	3626	100
765	D87446	Homo sapiens	Similar tc a C.elegans protein encoded in cosmid C27F2 (U40419)	568	38
766	AL023828	Caenorhabditis elegans	Y17G7B.14	200	27
767	Y82777	Homo sapiens	Human chordin related protein (Clone dw665_4).	2551	99
768	X92475	Homo sapiens	ITBA1	1429	100
769	Y42752	Homo sapiens	Human calcium binding protein 3 (CaBP-3).	1426	100
770	X51416	Homo sapiens	hormone receptor hERR1 (AA 1-521)	2641	97
771	AJ006591	Homo sapiens	cysteine-rich protein	1793	100
772	A08695	Homo sapiens	rap2	935	100
773	Z12173	Homo sapiens	N-acetylglucosamine-6-sulphatase	2970	100
774	Y91950	Homo sapiens	Human cytoskeleton associated protein 5 (CYSKP-5).	565	43
776	AL023799	Homo sapiens	dJ322P7.1 (zinc finger)	855	56
777	AL023799	Homo sapiens	dJ322P7.1 (zinc finger)	855	56
778	G01880	Homo sapiens	Human secreted protein, SEQ ID NO: 5961.	849	98
779	AJ012590	Homo sapiens	glucose 1-dehydrogenase	4155	99
780	AL078582	Homo sapiens	dJ130E4.2 (KIAA0796)	1321	68
781	Z75955	Caenorhabditis elegans	similar to mitochondrial carrier protein	384	34
782	AL109965	Homo sapiens	dJ1121G12.2 (SCAN domain-containing 1 protein)	900	100
783	AF061262	Mus musculus	semaP cytoplasmic domain associated protein 2	1316	83
784	G03873	Homo sapiens	Human secreted protein, SEQ	649	95

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			ID NO: 7954.		
785	Y84441	Homo sapiens	Amino acid sequence of a human RNA-associated protein.	2074	100
786	Y00918	Homo sapiens	Human Rab protein, RABP-1, protein sequence.	1048	99
787	Z97029	Homo sapiens	ribonuclease HI large subunit	1548	99
788	AB035384	Homo sapiens	SRp25 nuclear protein	962	94
789	AF024631	Homo sapiens	ANG2	2644	100
790	AJ006710	Rattus norvegicus	phosphatidylinositol 3-kinase	4508	97
792	V00638	bacteriophage lambda	reading frame ea10	600	100
793	AF049103	Homo sapiens	Huntingtin interacting protein	819	100
795	Z26317	Homo sapiens	desmoglein 2	4810	99
796	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	5080	99
797	U15155	Gallus gallus	trypsinogen	372	37
798	U97189	Caenorhabditis elegans	strong similarity to thw P13/P14 family of kinases	227	28
799	AF112201	Homo sapiens	neuronal protein NP25	1053	100
800	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	958	63
801	AF267852	Homo sapiens	placental protein 13-like protein	743	99
802	AF208851	Homo sapiens	BM-009	766	80
803	Z81097	Caenorhabditis elegans	Similarity to Human retinoblastoma-binding protein RBAP46 yk662d12.5 comes from this gene	152	27
804	G02113	Homo sapiens	Human secreted protein, SEQ ID NO: 6194.	496	98
805	AL121673	Homo sapiens	bA305P22.1 (novel protein)	1160	100
806	AC013483	Arabidopsis thaliana	putative GTPase activator protein	264	30
807	AC013483	Arabidopsis thaliana	putative GTPase activator protein	264	30
808	AB013885	Homo sapiens	beta-ureidopropionase	1494	100
809	AF078842	Homo sapiens	HOTTL protein	1581	99
810	AF161421	Homo sapiens	HSPC303	2134	96
811	AF261689	Homo sapiens	DNA polymerase epsilon p17 subunit	734	100
812	Z74029	Caenorhabditis elegans	Similarity to C.elegans alcohol dehydrogenase comes from this gene	610	71
813	Z73497	Homo sapiens	cU240C2.2 (Core histone H2A/H2B/H3/H4)	324	100
814	W87689	Homo sapiens	Human HTXFT19 polypeptide.	1484	99
815	X16282	Homo sapiens	zinc finger protein (217 AA) (1 is 2nd base in codon)	1109	99
816	Z92539	Mycobacterium tuberculosis	pth	300	36
818	AB030483	Mus musculus	B9	197	27
819	AL117555	Homo sapiens	hypothetical protein	321	94
820	AC005328	Homo sapiens	R26660_2, partial CDS	865	97
821	G03951	Homo sapiens	Human secreted protein, SEQ ID NO: 8032.	700	99
822	L34807	Musca domestica	transposase	174	20
823	G02928	Homo sapiens	Human secreted protein, SEQ ID NO: 7009.	558	78
824	Z99531	Schizosaccha	caffeine-induced death	184	29

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		romyces pombe	protein 1		
825	AJ006692	Homo sapiens	ultra high sulfer keratin	693	68
826	U23037	Oryctolagus cuniculus	eIF-2Bepsilon	3406	90
827	G03412	Homo sapiens	Human secreted protein, SEQ ID NO: 7493.	464	100
828	Y30827	Homo sapiens	Human secreted protein encoded from gene 17.	113	44
829	Y32199	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2022379.	1012	100
830	W78279	Homo sapiens	Fragment of human secreted protein encoded by gene 33.	1264	99
832	AB011542	Homo sapiens	MEGF9	2097	100
833	G02639	Homo sapiens	Human secreted protein, SEQ ID NO: 6720.	223	70
834	AF119664	Homo sapiens	transcriptional regulator protein HCNGP	1574	100
835	AF119664	Homo sapiens	transcriptional regulator protein HCNGP	1144	89
836	AF119664	Homo sapiens	transcriptional regulator protein HCNGP	1448	94
837	X12517	Homo sapiens	C protein (AA 1-159)	918	100
838	U32865	Drosophila melanogaster	linotte protein	164	24
839	AF067730	Homo sapiens	TLS-associated protein TASR-2	631	56
840	U27831	Homo sapiens	striatum-enriched phosphatase	2840	98
841	AF286366	Homo sapiens	CamKI-like protein kinase	1796	100
842	G02309	Homo sapiens	Human secreted protein, SEQ ID NO: 6390.	278	98
843	AE003615	Drosophila melanogaster	ade3 gene product	113	48
844	G01350	Homo sapiens	Human secreted protein, SEQ ID NO: 5431.	629	100
845	U27838	Mus musculus	glycosyl-phosphatidyl-inositol-anchored protein homolog	3305	96
847	Y87788	Homo sapiens	Human RBP-26 protein.	2026	100
848	AF164794	Homo sapiens	Diff33 protein homolog	2398	100
849	U41315	Homo sapiens	ZNF127-Xp	2458	93
850	AF192784	Homo sapiens	makorin 1	2062	97
851	Y58628	Homo sapiens	Protein regulating gene expression PRGE-21.	1548	100
852	Z22968	Homo sapiens	M130 antigen	6205	100
853	Z22971	Homo sapiens	M130 antigen extracellular variant	6380	100
854	G03362	Homo sapiens	Human secreted protein, SEQ ID NO: 7443.	330	96
855	G03362	Homo sapiens	Human secreted protein, SEQ ID NO: 7443.	203	100
856	AF285118	Homo sapiens	CGI-203	452	100
857	AC006069	Arabidopsis thaliana	putative cleavage and polyadenylation specifity factor	1383	55
858	AL021546	Homo sapiens	Cytochrome C Oxidase Polypeptide VIa-liver precursor (EC 1.9.3.1)	593	100
859	L02956	Xenopus laevis	ribonucleoprotein	1664	85
860	AF201947	Homo sapiens	MEK binding partner 1	616	100
861	L31783	Mus musculus	uridine kinase	1266	92
862	AF161472	Homo sapiens	HSPC123	602	73
863	Z49068	Caenorhabditis elegans	mitochondrial carrier protein	370	43
864	AF154108	Homo sapiens	tumor necrosis factor type 1	3559	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			receptor associated protein		
865	AE001530	Helicobacter pylori J99	putative	230	32
866	X57807	Homo sapiens	immunoglobulin lambda light chain	699	91
867	AL031673	Homo sapiens	dJ694B14.1 (PUTATIVE novel KRAB box protein with 18 C2H2 type Zinc finger domains)	4066	99
868	Y11652	Homo sapiens	phosphate cyclase	238	100
869	AF192968	Homo sapiens	high-glucose-regulated protein 8	3041	99
870	AB020648	Homo sapiens	KIAA0841 protein	3237	99
871	AL031427	Homo sapiens	dJ167A19.1 (novel protein)	1608	100
872	AF151534	Homo sapiens	core histone macroH2A2.2	1866	100
873	AL021331	Homo sapiens	dJ366N23.1 (putative C. elegans UNC-93 (protein 1, C46F11.1) LIKE protein)	1129	100
874	X14608	Homo sapiens	propionyl-CoA carboxylase	3579	100
875	AL117334	Homo sapiens	dJ687F11.1 (novel protein (part of translation of cDNA DKFZp434N061, Em:AL110249))	306	100
876	X79489	Saccharomyces cerevisiae	E-925 protein	446	35
877	Y53001	Homo sapiens	Human secreted protein clone dn834_1 protein sequence SEQ ID NO:8.	811	100
878	AF281064	Homo sapiens	CHMP1.5	957	100
879	X79417	Sus scrofa	40S ribosomal protein S12	687	100
880	AF001317	Saccharomyces cerevisiae	Soi1p	478	28
881	Y87275	Homo sapiens	Human signal peptide containing protein HSPP-52 SEQ ID NO:52.	2547	100
882	M14036	Homo sapiens	C1-inhibitor	598	77
883	AB041261	Homo sapiens	calcium-independent phospholipase A2	2903	100
884	AF020313	Mus musculus	proline-rich protein 48	999	84
885	Y10936	Homo sapiens	hypothetical protein	1104	99
886	AF073997	Mus musculus	myotubularin related protein 1	866	36
887	Y57893	Homo sapiens	Human transmembrane protein HTMPN-17.	1099	94
888	AL117635	Homo sapiens	hypothetical protein	929	99
889	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2046	99
890	Y36031	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 416.	583	100
891	Y36031	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 416.	192	57
892	AF237631	Homo sapiens	ubiquitous tropomodulin U-Tmod	1798	100
893	AF090929	Homo sapiens	PRO0477p	653	99
894	AL031228	Homo sapiens	dJ1033B10.2 (WD40 protein BING4 (similar to S. cerevisiae YER082C, M. sexta MNG10 and C. elegans F28D1.1)	3196	100
895	AL031228	Homo sapiens	dJ1033B10.2 (WD40 protein BING4 (similar to S. cerevisiae YER082C, M. sexta MNG10 and C. elegans F28D1.1)	2825	96
896	AF171102	Homo sapiens	retinal degeneration B beta	1302	95
897	AE003551	Drosophila melanogaster	CG18176 gene product	633	33

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
898	AJ237946	Homo sapiens	DEAD Box Protein 5	2443	100
899	Z97184	Homo sapiens	FKE2	624	100
900	Z97184	Homo sapiens	FKE2	409	98
901	AJ245587	Homo sapiens	Kruppel-type zinc finger	1942	100
902	AF091034	Homo sapiens	GTP-binding protein RAB22A	1011	100
903	R95953	Homo sapiens	Eukaryotic cell growth inhibiting factor.	414	96
904	L04733	Homo sapiens	kinesin light chain	1936	72
905	AE003540	Drosophila melanogaster	CG10984 gene product	446	33
906	M55542	Homo sapiens	guanylate binding protein isoform I	2993	98
907	M55542	Homo sapiens	guanylate binding protein isoform I	2901	96
908	W84085	Homo sapiens	Human membrane fusion protein WDProl.	1889	100
909	AF168676	Homo sapiens	TNF intracellular domain-interacting protein	647	100
910	AB029150	Homo sapiens	KRAB zinc finger protein HFB101L	2196	100
911	G02871	Homo sapiens	Human secreted protein, SEQ ID NO: 6952.	521	100
912	G03162	Homo sapiens	Human secreted protein, SEQ ID NO: 7243.	387	87
913	AJ243721	Homo sapiens] >Y92508 Y92508 13-APR-2000 06-OCT-1998 Human OXRE-5. [Homo sapiens	dTDP-4-keto-6-deoxy-D-glucose 4-reductase	1710	100
914	U24189	Caenorhabditis elegans	hypothetical protein 1207-1; Method: conceptual translation supplied by authors	244	41
915	Y02591	Homo sapiens	A human progesterone receptor complex p23-like protein.	843	99
916	AE000984	Archaeoglobus fulgidus	dinitrogenase reductase activating glycohydrolase (draG)	171	26
918	M23159	Cricetus cricetus	DHFR-coamplified protein	163	30
919	L12018	Caenorhabditis elegans	putative	1232	41
920	AF102177	Homo sapiens	tumor antigen SLP-8p	1260	97
921	AL096712	Homo sapiens	dJ744I24.2 (similar to a novel human gene mapping to Activator)	1017	78
922	AL161495	Arabidopsis thaliana	putative WD-repeat protein	866	42
923	AL161495	Arabidopsis thaliana	putative WD-repeat protein	442	36
924	U97001	Caenorhabditis elegans	similar to Schizosaccharomyces pombe	605	51
925	X71978	Mus musculus	Fif	1503	95
926	M92288	Drosophila melanogaster	beta-spectrin	290	51
927	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
928	Y22499	Homo sapiens	Human secreted protein sequence clone mh703_1.	2249	100
930	AJ224326	Homo sapiens	ribulose-5-phosphate-epimerase	912	100
931	U28991	Caenorhabditis	coded for by C. elegans cDNA	660	55

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		is elegans	cm21c7		
932	AL080065	Homo sapiens	hypothetical protein	210	25
933	G01884	Homo sapiens	Human secreted protein, SEQ ID NO: 5965.	767	98
934	AJ276485	Homo sapiens	integral membrane transporter protein	1200	100
935	AL035681	Homo sapiens	dJ756G23.3 (novel protein similar to drosophila transcriptional repressor)	1142	80
936	AB026808	Mus musculus	synaptotagmin XI	2142	95
937	AB015345	Homo sapiens	HRIHFB2216	2601	99
938	X65724	Homo sapiens	ORF2	498	100
939	W89024	Homo sapiens	Polypeptide fragment encoded by gene 156.	1487	100
940	G04047	Homo sapiens	Human secreted protein, SEQ ID NO: 8128.	117	100
941	AF094583	Homo sapiens	putative HIV-1 infection related protein	452	100
942	AC024200	Caenorhabditis elegans	contains similarity to several zinc finger proteins but not to the zinc finger domains	350	69
943	AF129756	Homo sapiens	G5c	273	100
944	M23765	Rattus norvegicus	alpha-tropomyosin	133	96
945	AC009917	Arabidopsis thaliana	Contains similarity to	583	47
946	AF223468	Homo sapiens	AD021 protein	551	44
947	AF055473	Homo sapiens	GAGE-8	273	51
948	X75756	Homo sapiens	protein kinase C mu	2019	68
949	AF143956	Mus musculus	coronin-2	2300	93
950	Y36729	Homo sapiens	Human PG1 protein sequence.	1861	99
951	W49041	Homo sapiens	Human low density lipoprotein binding protein LBP-2.	282	67
952	AB016881	Arabidopsis thaliana	gene_id:MXC17.7-	203	46
953	Y01785	Homo sapiens	Human ubiquitin-conjugating enzyme >Y25341 Y25341 01-JUL-1999 12-AUG-1998 Human NCE-2 protein.	365	100
954	AF145615	Drosophila melanogaster	BcDNA.GH03377	823	46
955	U09410	Homo sapiens	zinc finger protein ZNF131	2483	99
956	U09410	Homo sapiens	zinc finger protein ZNF131	1853	99
957	AF195623	Homo sapiens	cholinephosphotransferase 1 alpha	2126	99
958	X94917	Drosophila melanogaster	head-elevated expression in 0.9 kb	155	32
959	U54807	Rattus norvegicus	GTP-binding protein	1167	97
960	AF058807	Bos taurus	GTP-binding protein rah	606	97
961	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	471	100
962	AF078850	Homo sapiens	steroid dehydrogenase homolog	583	40
963	AP001754	Homo sapiens	transient receptor potential-related channel 7, a novel putative Ca2+ channel protein	317	30
964	AL035419	Homo sapiens	dJ1100H13.1 (putative novel protein)	1129	100
965	X61381	Rattus rattus	interferon-induced protein	202	46
966	D38169	Homo sapiens	inositol 1,4,5-trisphosphate 3-kinase isoenzyme	3278	100
967	AL031432	Homo sapiens	dJ465N24.2.1 (PUTATIVE novel protein) (isoform 1)	893	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
968	U79275	Homo sapiens	unknown	611	100
969	AJ011306	Homo sapiens	guanine nucleotide exchange factor (long isoform)	2752	99
970	AF281134	Homo sapiens	exosome component Rrp46	1186	100
971	U53336	Caenorhabditis elegans	weak similarity over a short region to myosin heavy chain	536	23
972	AC018749	Leishmania major	L8840.12	589	53
973	AF188504	Mus musculus	LNV	544	85
974	U25801	Homo sapiens	Tax1 binding protein	852	98
975	AF049523	Homo sapiens	huntingtin-interacting protein HYPA/FBP11	1390	97
976	AF161530	Homo sapiens	HSPC182	1040	100
977	G04020	Homo sapiens	Human secreted protein, SEQ ID NO: 8101.	626	100
978	AF164797	Homo sapiens	ribosomal protein L17 isolog	908	100
979	U94991	Xenopus laevis	transcription factor XLMO1	795	97
980	S73775	Homo sapiens	calmitine; calsequestrine	2029	100
981	Y94888	Homo sapiens	Human protein clone HP01462.	2501	100
982	AJ243191	Homo sapiens	heat shock protein	827	96
983	X65020	Bos taurus	PSST subunit of the NADH: ubiquinone oxidoreductase complex	964	85
984	AJ249207	Rhodococcus sp. AD45	putative racemase	351	43
985	Z30093	Homo sapiens	basic transcription factor 2, 35 kD subunit	1576	99
986	AB030835	Homo sapiens	contains two glutamine rich domains, three zinc-finger domains, and matrin 3 homologous domain 3 (MH3)	4697	99
987	AF227258	Bos taurus	RPGR-interacting protein-1	1262	38
988	AL022238	Homo sapiens	dJ1042K10.2 (supported by GENSCAN, FGENES and GENEWISE)	4048	99
989	AL022238	Homo sapiens	dJ1042K10.2 (supported by GENSCAN, FGENES and GENEWISE)	2321	99
990	AF161426	Homo sapiens	HSPC308	448	92
991	AF161426	Homo sapiens	HSPC308	448	92
992	AF161426	Homo sapiens	HSPC308	453	92
993	AL023859	Schizosaccharomyces pombe	trna-splicing endonuclease subunit	172	42
994	AL049631	Homo sapiens	dJ513M9.1 (novel Homeobox domain protein)	241	47
995	AC005253	Homo sapiens	R26445_1	902	100
996	AF265206	Homo sapiens	MOG1 isoform A	974	100
997	AJ248285	Pyrococcus abyssi	sarcosine oxidase, subunit beta (soxB)	195	28
998	AE003641	Drosophila melanogaster	BG:DS00941.3 gene product	218	58
999	W69343	Homo sapiens	Secreted protein of clone CR930_1.	1340	98
1000	AY007135	Homo sapiens	similar to bovine ADP/ATP translocase T1 mRNA with GenBank Accession Number M24102.1	1543	100
1001	Y73381	Homo sapiens	HTRM clone 1877278 protein sequence.	1668	100
1002	AF208844	Homo sapiens	BM-002	428	100
1003	AE004944	Pseudomonas aeruginosa	hypothetical protein	134	35
1004	AL031431	Homo sapiens	dJ462023.2 (novel protein)	2058	100
1005	S45367	Canis familiaris	centractin	1949	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1006	S45367	Canis familiaris	centractin	1315	98
1007	AB022158	Mus musculus	chaperonin containing TCP-1 epsilon subunit	2649	96
1008	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	1282	97
1009	AB011414	Homo sapiens	Kruppel-type zinc finger protein	1671	58
1010	Z68218	Caenorhabditis elegans	K01H12.1	269	67
1011	AB011414	Homo sapiens	Kruppel-type zinc finger protein	1671	58
1012	Z14000	Homo sapiens	RING1	2017	100
1013	G02841	Homo sapiens	Human secreted protein, SEQ ID NO: 6922.	332	93
1014	AF145659	Drosophila melanogaster	BcDNA.GH10333	1244	52
1015	Y02860	Homo sapiens	Fragment of human secreted protein encoded by gene 65.	664	67
1016	Y02591	Homo sapiens	A human progesterone receptor complex p23-like protein.	772	97
1017	Y99448	Homo sapiens	Human PRO1759 (UNQ832) amino acid sequence SEQ ID NO:374.	2323	100
1018	X67250	Rattus norvegicus	n-chimaerin	1710	97
1019	AF183417	Homo sapiens	microtubule-associated proteins 1A/1B light chain 3	631	100
1020	AF164795	Homo sapiens	sex-regulated protein janus-a	674	100
1021	AF190625	Coturnix coturnix	qdgl-1	638	96
1022	AL133363	Arabidopsis thaliana	putative protein	155	37
1023	AB034912	Homo sapiens	WD-repeat like sequence	2483	100
1024	AY007091	Homo sapiens	similar to Homo sapiens mammalian inositol hexakisphosphate kinase 2 (IP6K2) mRNA with Ge	2243	100
1025	X69910	Homo sapiens	P63 protein	2958	99
1026	U80736	Homo sapiens	CAGF9	1657	100
1027	AB029333	Halocynthia roretzi	HrPET-1	1048	54
1028	AB032931	Homo sapiens	ubiquitin-conjugating enzyme isolog	1045	100
1029	G01797	Homo sapiens	Human secreted protein, SEQ ID NO: 5878.	749	98
1030	G01797	Homo sapiens	Human secreted protein, SEQ ID NO: 5878.	749	98
1031	AF193795	Homo sapiens	vacuolar sorting protein VPS29/PEP11	960	100
1032	AJ222968	Mus musculus	L-periaxin	120	30
1033	Z81317	Schizosaccharomyces pombe	DNA2-NAM7 helicase family protein	685	31
1034	Y41519	Homo sapiens	Fragment of human secreted protein encoded by gene 75.	1321	99
1035	AJ276004	Mus musculus	Paxneb protein	1709	77
1036	AF025459	Caenorhabditis elegans	H14A12.3 gene product	190	30
1037	U37251	Homo sapiens	Description: KRAB zinc finger protein; this is a splicing supplied by author	196	43
1038	W74580	Homo sapiens	Human membrane protein BA0306.	1921	97
1039	U88173	Caenorhabditis elegans	weak similarity to Arabidopsis thaliana ubiquitin-like protein 8	331	80

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1040	AF290204	Homo sapiens	blood group carrier molecule DOK1	1637	99
1041	Y96730	Homo sapiens	PRO539, a Costal-2 homologue.	162	22
1042	AF140683	Mus musculus	F-box protein FWD2	2397	98
1043	AF151023	Homo sapiens	HSPC189	1104	100
1044	AF181631	Drosophila melanogaster	BcDNA.GH04929	204	37
1045	Y77985	Homo sapiens	Human collectin amino acid sequence.	1940	100
1046	AJ243972	Homo sapiens	6-phosphogluconolactonase	1317	100
1047	AB035863	Homo sapiens	ATP specific succinyl CoA synthetase beta subunit precursor	2324	99
1048	AL034550	Homo sapiens	dJ1184F4.2 (novel protein similar to nucleolar protein 4 (NOL4) (NOLP))	981	92
1049	AF163825	Homo sapiens	pre-B lymphocyte protein 3	634	100
1050	AF201949	Homo sapiens	60S ribosomal protein L30 isolog	868	100
1051	AF190624	Mus musculus	mdgl-1	236	85
1052	AE003529	Drosophila melanogaster	CG6151 gene product	160	44
1053	G01191	Homo sapiens	Human secreted protein, SEQ ID NO: 5272.	646	98
1054	AL162756	Neisseria meningitidis	Glu-tRNA(Gln) amidotransferase subunit A	682	44
1055	AF181856	Rattus norvegicus	tRNA selenocysteine associated protein	1525	99
1056	U89649	Chlamydomonas reinhardtii	Mr19,000 outer arm dynein light chain	244	34
1057	AF159141	Homo sapiens	breast cancer metastasis-suppressor 1	663	53
1058	AF230929	Homo sapiens	keratinocyte annexin-like protein pemphaxin	1710	99
1059	AJ270952	Homo sapiens	putative membrane protein	1363	100
1060	AF224263	Heterodontus francisci	HoxD8	742	83
1061	X63417	Homo sapiens	IRLB	1037	100
1062	AL079345	Streptomyces coelicolor A3(2)	hypothetical protein	143	27
1063	Y71112	Homo sapiens	Human Hydrolase protein-10 (HYDRL-10).	2547	100
1064	AF263614	Homo sapiens	acetyl-CoA synthetase	3493	99
1065	Y13356	Homo sapiens	Amino acid sequence of protein PRO221.	1363	100
1066	AC006153	Homo sapiens	similar to Aquifex aeolicus GTP-binding protein; similar to AE000771 (PID:g2984292)	662	98
1067	Y18930	Sulfolobus solfataricus	hypothetical protein	162	29
1068	R65969	Homo sapiens T98G	Glioblastoma-derived polypeptide.	887	100
1069	Y07964	Homo sapiens	Human secreted protein fragment	863	96
1070	AF177476	Rattus norvegicus	CDK5 activator-binding protein	1995	86
1071	AF245505	Homo sapiens	adlican	3109	99
1072	U92794	Mus musculus	alpha glucosidase II, beta subunit	147	36
1073	G03889	Homo sapiens	Human secreted protein, SEQ ID NO: 7970.	698	98
1074	U15779	Homo sapiens	p70	380	28
1075	Y13392	Homo sapiens	Amino acid sequence of	1271	91

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			protein PRO328.		
1076	AF161457	Homo sapiens	HSPC339	571	100
1077	Y79509	Homo sapiens	Human carbohydrate-associated protein CRBAP-5.	2151	98
1078	AF223466	Homo sapiens	HT015 protein	831	66
1079	AL132965	Arabidopsis thaliana	putative WD-40 repeat-protein	286	29
1080	AB024937	Homo sapiens	LUNX	1284	100
1081	Y14768	Homo sapiens	V-ATPase G-subunit like protein	579	100
1082	AF016416	Caenorhabditis elegans	F29A7.4 gene product	141	31
1083	L13291	Homo sapiens	ADP-ribosylarginine hydrolase	802	45
1084	AB041541	Mus musculus	unnamed protein product	151	44
1085	G01922	Homo sapiens	Human secreted protein, SEQ ID NO: 6003.	202	97
1086	AB030814	Homo sapiens	H-REV107 protein homolog	833	100
1087	AF151638	Homo sapiens	phosphatidylcholine transfer protein	1142	100
1088	Y84432	Homo sapiens	Amino acid sequence of a human RNA-associated protein.	2783	100
1089	Y94867	Homo sapiens	Human protein clone HP10563.	613	100
1090	AK023982	Homo sapiens	unnamed protein product	130	49
1091	AB041586	Mus musculus	unnamed protein product	1103	81
1092	Y71277	Homo sapiens	Human Zlipo3 protein.	606	100
1093	U34973	Mus musculus	protein tyrosine phosphatase-like	1131	95
1094	Y66677	Homo sapiens	Membrane-bound protein PRO828.	522	56
1095	Y87276	Homo sapiens	Human signal peptide containing protein HSPP-53 SEQ ID NO:53.	1029	99
1096	Y87276	Homo sapiens	Human signal peptide containing protein HSPP-53 SEQ ID NO:53.	863	98
1097	AF161455	Homo sapiens	HSPC337	742	98
1098	U80029	Caenorhabditis elegans	similar to thioredoxin	242	39
1099	AJ005866	Homo sapiens	Sqv-7-like protein	1321	99
1100	AJ005866	Homo sapiens	Sqv-7-like protein	1118	99
1101	AJ005866	Homo sapiens	Sqv-7-like protein	891	99
1102	AJ005866	Homo sapiens	Sqv-7-like protein	1016	99
1103	AL110244	Homo sapiens	hypothetical protein	299	31
1104	AF242194	Drosophila melanogaster	brakeless-B	147	52
1105	AL031010	Homo sapiens	dJ422F24.1 (PUTATIVE novel protein similar to C. elegans C02C2.5)	968	100
1106	U28016	Mus musculus	parathion hydrolase (phosphotriesterase)-related protein	1624	87
1107	AJ278150	Homo sapiens	putative lipid kinase	2207	99
1108	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	495	98
1109	AF217287	Drosophila melanogaster	G protein RhoBTB	834	54
1110	Y28921	Homo sapiens	Human regulatory protein HRGP-7.	941	48
1111	Y28921	Homo sapiens	Human regulatory protein HRGP-7.	1331	51
1112	AF176704	Homo sapiens	F-box protein FBX9	2027	99
1113	AF182076	Homo sapiens	glioma tumor suppressor candidate region protein 2	2418	100
1114	G04039	Homo sapiens	Human secreted protein, SEQ	475	96

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			ID NO: 8120.		
1115	AF229439	Mus musculus	zinc finger protein 289	1697	91
1116	L40357	Homo sapiens	thyroid receptor interactor	509	100
1117	L40357	Homo sapiens	thyroid receptor interactor	404	85
1118	A12155	Homo sapiens	Human X5L cDNA.	1673	100
1119	AL161542	Arabidopsis thaliana	isomerase like protein	607	53
1120	AL023754	Homo sapiens	dJ272L16.1 (Rat Ca2+/Calmodulin dependent Protein Kinase LIKE protein)	2341	98
1121	Y57901	Homo sapiens	Human transmembrane protein ETMPN-25.	321	36
1122	Z14122	Xenopus laevis	XLCL2	455	77
1123	AF225418	Homo sapiens	lipase	1531	97
1124	Y06518	Homo sapiens	Zen GTPase interacting protein ZIP.	3227	100
1125	AL035690	Homo sapiens	dJ202I21.1 (novel protein)	952	100
1126	AJ000217	Homo sapiens	CLIC2	1286	99
1127	AB030505	Mus musculus	UBE-1c2	1069	79
1128	Y73375	Homo sapiens	HTRM clone 1427838 protein sequence.	874	100
1129	Y78941	Homo sapiens	Cyclophilin-type peptidyl prolyl cis/trans isomerase amino acid sequence.	877	100
1130	AL023553	Homo sapiens	dJ347H13.4 (novel protein)	557	100
1131	Y91945	Homo sapiens	Human chaperone protein 6 (HCHP-6).	1408	100
1132	Z68197	Schizosaccharomyces pombe	putative nuclear pore protein	596	39
1133	Z68197	Schizosaccharomyces pombe	putative nuclear pore protein	389	35
1134	AF180681	Homo sapiens	guanine nucleotide exchange factor	3597	100
1135	AF079765	Mus musculus	enhancer of polycomb	264	41
1136	M62419	Mus musculus	clathrin-associated protein	2189	99
1137	AJ006219	Drosophila melanogaster	clathrin-associated protein	1254	78
1138	Y76218	Homo sapiens	Human secreted protein encoded by gene 95.	440	98
1139	W88104	Homo sapiens	A Rab protein designated HRABS-2.	1065	99
1140	Y13401	Homo sapiens	Amino acid sequence of protein PRO339.	3979	98
1141	W85026	Chimeric - Homo sapiens	Green fluorescent protein-Zap70 fusion product.	3309	100
1142	Y13402	Homo sapiens	Amino acid sequence of protein PRO310.	1694	99
1143	G03875	Homo sapiens	Human secreted protein, SEQ ID NO: 7956.	660	99
1144	Y12917	Homo sapiens	Amino acid sequence of a human secreted peptide.	750	98
1145	Y12917	Homo sapiens	Amino acid sequence of a human secreted peptide.	1096	100
1146	AL022157	Homo sapiens	SPIN (SPINDLIN HOMOLOG (PROTEIN DXF34))	1233	100
1147	AL022157	Homo sapiens	SPIN (SPINDLIN HOMOLOG (PROTEIN DXF34))	1233	100
1148	G02548	Homo sapiens	Human secreted protein, SEQ ID NO: 6629.	370	98
1149	Y73338	Homo sapiens	HTRM clone 2019742 protein sequence.	1492	100
1150	W74841	Homo sapiens	Human secreted protein encoded by gene 113 clone	228	55

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			HEAAR60.		
1151	AF044201	Rattus norvegicus	neural membrane protein 35; NMP35	1570	92
1152	AF156774	Homo sapiens	lysophosphatidic acid acyltransferase-gammal	1855	99
1153	AL118501	Homo sapiens	dJ1191N16.1 (A novel protein (translation of the cDNA DKFZp566A0946, Em:AL050069))	872	64
1154	AF131852	Homo sapiens	Unknown	473	100
1155	Y41705	Homo sapiens	Human PRO352 protein sequence.	1381	97
1156	G04036	Homo sapiens	Human secreted protein, SEQ ID NO: 8117.	607	99
1157	AF112444	Lupinus luteus	L-asparaginase	287	43
1158	AF151848	Homo sapiens	CGI-90 protein	232	32
1159	AJ272267	Homo sapiens	choline dehydrogenase	2449	100
1160	AB001773	Ciona savignyi	PEM-6	196	33
1161	Y87330	Homo sapiens	Human signal peptide containing protein HSPP-107 SEQ ID NO:107.	746	83
1162	Y87330	Homo sapiens	Human signal peptide containing protein HSPP-107 SEQ ID NO:107.	746	83
1163	AF113534	Homo sapiens	HP1-BP74 protein	2723	96
1164	AF232226	Danio rerio	Deddl	191	41
1165	AL118501	Homo sapiens	dJ1191N16.1 (A novel protein (translation of the cDNA DKFZp566A0946, Em:AL050069))	1051	71
1166	AL118501	Homo sapiens	dJ1191N16.1 (A novel protein (translation of the cDNA DKFZp566A0946, Em:AL050069))	945	76
1167	AF187733	Homo sapiens	syntaphilin	831	42
1168	AB019435	Homo sapiens	phospholipase	951	55
1169	AF064604	Homo sapiens	KE03 protein	324	33
1170	Y01164	Homo sapiens	Polypeptide fragment encoded by gene 6.	1191	100
1171	L03188	Saccharomyces cerevisiae	putative	180	22
1172	AF113751	Mus musculus	nuclear pore membrane glycoprotein POM210	3941	81
1173	AJ245417	Homo sapiens	G5b protein	794	100
1174	AL022238	Homo sapiens	dJ1042K10.3 (novel protein)	1285	100
1175	U41278	Caenorhabditis elegans	F33G12.3 gene product	332	28
1176	M35617	Homo sapiens	T-cell receptor V-alpha-J-alpha region	284	83
1177	AC012680	Arabidopsis thaliana	putative protein phosphatase 2C; 55455-56414	209	37
1178	G01345	Homo sapiens	Human secreted protein, SEQ ID NO: 5426.	692	99
1179	AL096767	Homo sapiens	dJ579N16.3 (novel protein similar to worm, Arabidopsis and pine proteins)	1342	100
1180	AF039716	Caenorhabditis elegans	similar to ATP synthase B chain	496	55
1181	Y11710	Homo sapiens	collagen type XIV	1048	97
1182	X82240	Homo sapiens] >R94974 R94974 09-MAY-1996 27-OCT-1994 Human TCL-1 polypeptide.	T cell leukemia/lymphoma 1	617	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		[Homo sapiens			
1183	U42841	Caenorhabditis elegans	short region of weak similarity to collagen	161	33
1185	AJ131613	Homo sapiens	dicarboxylate carrier protein	1470	99
1186	L27645	Danio rerio	growth-associated protein	130	36
1187	Y02738	Homo sapiens	Human secreted protein encoded by gene 89 clone HLHFP03.	636	100
1188	AF217544	Xenopus laevis	ornithine decarboxylase-2	1459	60
1189	AL136307	Homo sapiens	dJ380B8.2 (Neuritin, a protein which promotes neurite outgrowth)	182	33
1190	X89602	Homo sapiens	rTSbeta	197	100
1191	U32828	Haemophilus influenzae Rd	ribosomal protein S6 modification protein (rimK)	268	31
1192	AF154831	Rattus norvegicus	PV-1	1403	60
1193	Y50926	Homo sapiens	Human fetal brain cDNA clone vc16_1 derived protein.	918	100
1194	AF026530	Rattus norvegicus	stathmin-like-protein splice variant RB3''	1093	97
1195	U35244	Rattus norvegicus	vacuolar protein sorting homolog r-vps33a	2981	96
1196	Y70470	Homo sapiens	Human p53 target molecule, PRG3 protein.	1680	100
1197	AF157318	Homo sapiens	AD-017 protein	912	47
1198	AF125443	Caenorhabditis elegans	contains similarity to S. pombe phosphatidyl synthase (GB:Z28295)	460	39
1199	AF201934	Homo sapiens	DC12	1649	88
1200	AL031775	Homo sapiens	dJ30M3.3 (novel protein similar to C. elegans Y63D3A.4)	1902	100
1201	M21103	Ovis aries	BIIIB4 high-sulfur keratin	484	82
1202	Z85986	Homo sapiens	dJ108K11.3 (similar to yeast suppressor protein SRP40)	1143	75
1203	U18762	Rattus norvegicus	retinol dehydrogenase type I	890	52
1204	U35730	Mus musculus	jerky	2235	76
1205	AB002327	Homo sapiens	KIAA0329	151	24
1206	AB019233	Arabidopsis thaliana	ubiquinone/menaquinone biosynthesis methyltransferase-like	762	56
1207	AL136307	Homo sapiens	dJ380B8.2 (Neuritin, a protein which promotes neurite outgrowth)	742	100
1208	AF207989	Homo sapiens	orphan G-protein coupled receptor	2326	100
1209	Z97630	Homo sapiens	dJ466N1.4 (novel protein similar to ANK3 (ankyrin 3, node of Ranvier (ankyrin G)))	181	44
1210	U21549	Mus musculus	Ac39/physophilin	1280	68
1211	Y27700	Homo sapiens	Human secreted protein encoded by gene No. 12.	1267	100
1212	AF117814	Mus musculus	odd-skipped related 1 protein	945	66
1213	AF277233	Naegleria fowleri	calcineurin B	222	39
1214	D14849	Mus musculus	meiosis-specific nuclear structural protein 1	1950	77
1215	G03022	Homo sapiens	Human secreted protein, SEQ ID NO: 7103.	590	100
1216	Z72510	Caenorhabdit	similarity to yeast UTR3	634	49

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		is elegans	protein (Swiss Prot accession yk677h11.5 comes from this gene		
1217	Z49703	Saccharomyces cerevisiae	unknown	134	22
1218	AC013430	Arabidopsis thaliana	F3F9.18	199	29
1219	L10910	Homo sapiens	splicing factor	1026	71
1220	Z70750	Caenorhabditis elegans	similar to vanadate resistance protein transmembranous comes from this gene	965	58
1221	AL163815	Arabidopsis thaliana	putative protein	653	61
1222	AF155100	Homo sapiens	zinc finger protein KY-REN-21 antigen	2261	100
1223	J05071	Bos taurus	GTP-binding regulatory protein gamma-6 subunit	356	100
1224	Y73364	Homo sapiens	HTRM clone 2765991 protein sequence.	1169	99
1225	AL050170	Homo sapiens	hypothetical protein	714	100
1226	X64002	Homo sapiens	RAP74	2661	99
1227	X04085	Homo sapiens	catalase	2846	100
1228	AJ005620	Mus musculus	skeletal muscle-specific gene	1416	90
1229	AF045564	Rattus norvegicus	development-related protein	1715	93
1230	X97571	Mus musculus	HCMV-interacting protein	479	96
1231	L08239	Homo sapiens	located at OATL1	2274	100
1232	AF121863	Homo sapiens	sorting nexin 14	1964	100
1233	AF121863	Homo sapiens	sorting nexin 14	1203	84
1234	AC024805	Caenorhabditis elegans	contains similarity to TR:O04595	744	31
1235	AC006634	Caenorhabditis elegans	contains similarity to Saccharomyces cerevisiae probable membrane protein YLR418c (GB:U20162)	357	33
1236	Y18101	Mus musculus	macrophage actin-associated-tyrosine-phosphorylated protein	1559	87
1237	AB042646	Homo sapiens	TGIF2	1224	100
1238	AB026264	Homo sapiens	IMPACT	1694	100
1239	AB026264	Homo sapiens	IMPACT	1123	100
1240	G00429	Homo sapiens	Human secreted protein, SEQ ID NO: 4510.	324	100
1241	Y76144	Homo sapiens	Human secreted protein encoded by gene 21.	1363	53
1242	AL035602	Arabidopsis thaliana	putative protein	499	28
1243	X76483	Gallus gallus	Yes-associated protein (65kDa)	574	48
1244	AF220186	Homo sapiens	uncharacterized hypothalamus protein HT012	503	100
1245	AL021453	Homo sapiens	dJ821D11.3 (PUTATIVE protein)	856	100
1246	AJ276003	Homo sapiens	GAR1 protein	1216	100
1247	Y57910	Homo sapiens	Human transmembrane protein HTMPN-34.	1369	98
1248	AC004874	Homo sapiens	similar to N-acetylgalactosaminyltransferase; similar to Q07537 (PID:g1171989)	957	100
1249	AF199597	Homo sapiens	A-type potassium channel modulatory protein 1	1139	100
1250	Y13148	Rattus norvegicus	PAG608	1350	88
1251	M24852	Rattus norvegicus	neuron-specific protein PEP-19	124	46

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1252	AF146738	Rattus norvegicus	testis specific protein	771	83
1253	G02725	Homo sapiens	Human secreted protein, SEQ ID NO: 6806.	419	97
1254	W44375	Homo sapiens	Human ubiquitin-conjugating enzyme polypeptide.	1045	99
1255	AC006538	Homo sapiens	BC41195_1	831	78
1256	AB004316	Bos taurus	mitochondrial methionyl-tRNA transformylase	1556	88
1257	Z35094	Homo sapiens	SURF-2	1354	97
1258	Y13362	Homo sapiens	Amino acid sequence of protein PRO214.	2383	100
1259	AC006014	Homo sapiens	similar to RFP transforming protein; similar to P14373 (PID:g132517)	1299	100
1260	AC005099	Homo sapiens	match to AI222572 (NID:g3804775)	469	100
1261	V00507	Homo sapiens	coding sequence of DHFR (1 is 1st base in codon) (561 is 3rd base in codon)	984	100
1262	X15443	Rattus sp.	gamma-glutamyltranspeptidase (AA 1-568)	697	32
1263	AF173871	Mus musculus	neuronal PAS3	977	94
1264	AF178983	Homo sapiens	Ras-associated protein Rap1	433	97
1265	Y70473	Homo sapiens	Human cyclic nucleotide-associated protein-1 (CNAP-1).	2785	99
1266	Y41738	Homo sapiens	Human PRO541 protein sequence.	1622	100
1267	AF061346	Mus musculus	Edp1 protein	1077	64
1268	U97006	Caenorhabditis elegans	C13F10.4 gene product	154	23
1269	AF233582	Mus musculus	GTPase Rab37	942	95
1270	AF195951	Homo sapiens	signal recognition particle 68	3127	98
1271	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	1150	55
1272	AF201933	Homo sapiens	DC11	650	100
1273	AF201933	Homo sapiens	DC11	346	98
1274	AL021710	Arabidopsis thaliana	putative protein	348	49
1275	AC004449	Homo sapiens	R33683_3	556	100
1276	Y86295	Homo sapiens	Human secreted protein HL2AG87, SEQ ID NO:210.	1920	100
1277	Y71111	Homo sapiens	Human Hydrolase protein-9 (HYDRL-9).	1576	99
1278	S94421	Homo sapiens	T cell receptor eta-exon	478	100
1279	Y66695	Homo sapiens	Membrane-bound protein PRO1344.	1909	100
1280	AF161380	Homo sapiens	HSPC262	772	100
1281	Y48610	Homo sapiens	Human breast tumour-associated protein 71.	779	100
1282	AC015446	Arabidopsis thaliana	Similar to AIG1 protein	406	35
1283	AK024432	Homo sapiens	FLJ00022 protein	403	35
1284	W96153	Homo sapiens	Human FADD-interacting protein (FIP).	1825	81
1285	AJ001019	Homo sapiens	ring finger protein	1301	100
1286	AE003823	Drosophila melanogaster	CG13178 gene product	195	29
1287	AF178632	Homo sapiens	FEM-1-like death receptor binding protein	3261	100
1288	AC006033	Homo sapiens	similar to MLN 64; similar to I38027 (PID:g2135214)	1195	100
1289	AC006033	Homo sapiens	similar to MLN 64; similar to I38027 (PID:g2135214)	668	93
1290	AB023811	Homo sapiens	TU3A	351	54

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1291	Z73424	Caenorhabditis elegans	C44B9.1	235	36
1292	Y94871	Homo sapiens	Human protein clone HP02551.	1222	100
1293	AF190425	Homo sapiens	retinoblastoma-associated protein RAP140	489	29
1294	G03856	Homo sapiens	Human secreted protein, SEQ ID NO: 7937.	538	99
1295	AF133670	Mus musculus	ARL-6 interacting protein-2	367	51
1296	AJ249735	Homo sapiens	claudin-6	1142	100
1297	X57560	Escherichia coli	pspE protein	535	100
1298	AF169284	Homo sapiens	LIM and cysteine-rich domains protein 1	1997	100
1299	U41023	Caenorhabditis elegans	coded for by C. elegans cDNA yk61f1.3; coded for by C. yk109h8.5	324	29
1300	AB024523	Homo sapiens	basic kruppel like factor	1206	100
1301	X55989	Homo sapiens	eosinophil cationic-related protein	737	99
1302	AF007151	Homo sapiens	unknown	1481	100
1303	X52904	Escherichia coli	open reading frame (AA 1-65)	359	100
1304	U19577	Escherichia coli	galactonate dehydratase	242	93
1305	AF266508	Mus musculus	NELF protein	1409	97
1306	Y57901	Homo sapiens	Human transmembrane protein HTMPN-25.	932	100
1307	U58750	Caenorhabditis elegans	similar to the mitochondrial carrier family	365	54
1308	AF044774	Homo sapiens	breakpoint cluster region protein 2	2681	99
1309	AL078593	Homo sapiens	dJ210B1.1 (KIAA0680)	267	34
1310	X82693	Homo sapiens	E48 antigen	620	96
1311	Z82263	Caenorhabditis elegans	C47A4.1	283	35
1312	AF131218	Homo sapiens	chromosome 16 open reading frame 5	1493	100
1313	Y41763	Homo sapiens	Human PRO938 protein sequence.	1636	100
1314	AF196972	Homo sapiens	JM24 protein	2239	100
1315	AF053356	Homo sapiens	insulin receptor substrate like protein	228	97
1316	Y66695	Homo sapiens	Membrane-bound protein PRO1344.	1909	100
1317	AF153127	Gallus gallus	SAPK interacting protein	2442	89
1318	AF153127	Gallus gallus	SAPK interacting protein	1477	83
1319	AF153127	Gallus gallus	SAPK interacting protein	1651	86
1320	X56932	Homo sapiens	23 kD highly basic protein	1044	100
1321	AF174605	Homo sapiens] >Y83086 Y83086 09-MAR-2000 28-AUG-1998 F-box protein FBP-18. [Homo sapiens	F-box protein Fbx25	467	70
1322	M61732	Trypanosoma cruzi	neuraminidase	214	24
1323	Y17013	porcine endogenous	pol	304	64

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		retrovirus			
1324	AL138655	Arabidopsis thaliana	putative protein	1174	37
1325	AL138655	Arabidopsis thaliana	putative protein	946	35
1326	AL133215	Homo sapiens	bA108L7.2 (novel protein similar to rat tricarboxylate carrier)	1322	99
1327	AF161541	Homo sapiens	HSPC056	1357	99
1328	Y73346	Homo sapiens	HTRM clone 69699 protein sequence.	785	96
1329	L10910	Homo sapiens	splicing factor	912	82
1330	AF146568	Homo sapiens	MIL1 protein	1936	100
1331	W87772	Homo sapiens	Human serum glucocorticoid-regulated kinase (H-SGK2) polypeptide.	232	39
1332	Y41741	Homo sapiens	Human PRO704 protein sequence.	1860	100
1333	AF295096	Homo sapiens	zinc-finger protein ZBRK1	411	91
1334	Z82271	Caenorhabditis elegans	Similarity to Mouse kinensin-like protein KIF4 comes from this gene	578	44
1335	AE000810	Methanobacterium thermoautotrophicum	conserved protein	290	43
1336	Y68779	Homo sapiens	Amino acid sequence of a human phosphorylation effector PHSP-11.	1019	91
1337	AB027003	Mus musculus	protein phosphatase	378	84
1338	U64856	Caenorhabditis elegans	weak similarity to TPR domains	215	40
1339	AE001394	Plasmodium falciparum	protein of the YMR7 family	170	29
1340	X76717	Homo sapiens	MT-11 protein	204	89
1341	AC011914	Arabidopsis thaliana	putative mutT protein; 68398-67881	289	45
1342	AJ276171	Homo sapiens	ASPIC	2122	100
1343	AF187016	Homo sapiens	myosin regulatory light chain interacting protein MIR	2303	99
1344	AC006963	Homo sapiens	similar to Kelch proteins; similar to BAA77027 (PID:g4650844)	894	35
1345	AF257466	Homo sapiens	N-acetylneuraminic acid phosphate synthase	1880	99
1346	Y25896	Homo sapiens	Human secreted protein fragment encoded from gene 64.	1148	100
1347	AJ272073	Torpedo marmorata	male sterility protein 2-like protein	1664	58
1348	AF161548	Homo sapiens	HSPC063	1018	98
1349	W78128	Homo sapiens	Human secreted protein encoded by gene 3 clone HOSBI96.	1117	100
1351	G02144	Homo sapiens	Human secreted protein, SEQ ID NO: 6225.	418	100
1352	D90869	Escherichia coli	similar to	2047	100
1353	A12029	Homo sapiens	MRP-14	613	100
1354	AC005328	Homo sapiens	R26660_1, partial CDS	870	74
1355	AC024876	Caenorhabditis elegans	contains similarity to SW:RPB1_CRIGR	829	61
1356	AF077226	Homo sapiens	copine III	1876	64
1359	AF217188	Mus musculus	YIP1B	801	63
1360	AC074331	Homo sapiens	ZNF234	3869	100
1361	AL163279	Homo sapiens	homolog to cAMP response	5035	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			element binding and beta transducin family proteins		
1362	Z48475	Homo sapiens	glucokinase regulator	3160	99
1363	Z48475	Homo sapiens	glucokinase regulator	2682	97
1364	AF195764	Homo sapiens	megakaryocyte-enhanced gene transcript 1 protein; MEGT1 protein	2055	99
1365	AF116609	Homo sapiens	PRO0915	581	100
1366	AF116609	Homo sapiens	PRO0915	581	100
1367	AL117352	Homo sapiens	dJ876B10.3 (novel protein similar to C. elegans T19B10.6 (Tr:Q22557))	2581	99
1368	Y34124	Homo sapiens	Human potassium channel K+Hnov15.	1342	100
1369	AJ245621	Homo sapiens	CTL2 protein	3728	99
1370	AF008220	Bacillus subtilis	YtaG	429	45
1371	X05562	Homo sapiens	alpha-2 chain precursor (AA - 25 to 1018) (3416 is 2nd base in codon)	5908	99
1372	Z98048	Homo sapiens	dJ408N23.4 (novel DnaJ domain protein)	1296	99
1373	AF154415	Homo sapiens	FLASH	10253	100
1374	U20286	Rattus norvegicus	lamina associated polypeptide 1C	1567	69
1375	U53445	Homo sapiens	DOC1	1645	46
1376	AL117337	Homo sapiens	bA393J16.1 (zinc finger protein 33a (KOX 31))	250	60
1377	AC005328	Homo sapiens	R26660_1, partial CDS	1126	100
1378	U35113	Homo sapiens	metastasis-associated gene	1823	69
1379	L15313	Caenorhabditis elegans	putative	858	58
1380	Y25756	Homo sapiens	Human secreted protein encoded from gene 46.	1508	100
1381	AB037360	Homo sapiens	ANKHZN	5734	95
1382	AB037360	Homo sapiens	ANKHZN	959	97
1383	AF237676	Mus musculus	G beta-like protein GBL	1721	96
1384	AF237676	Mus musculus	G beta-like protein GBL	1043	70
1385	Y58793	Homo sapiens	Human calcium regulatory protein CaREG-1.	715	100
1386	AF212162	Homo sapiens	ninein	10369	99
1387	AL031685	Homo sapiens	dJ963K23.2 (novel protein)	337	33
1388	AC004890	Homo sapiens	similar to zinc finger proteins; similar to BAA24380 >W06316 W06316 03-OCT-1996 27-APR-1995 TRP-1 protein.	542	86
1389	AF187989	Homo sapiens	zinc finger protein ZNF223	2665	99
1390	AC035150	Homo sapiens	Zinc finger protein ZNF221	3459	100
1391	AF287894	Homo sapiens	PIST	1410	97
1392	AF282265	Homo sapiens	inner centromere protein INCENP	1794	99
1393	X90840	Homo sapiens	axonal transporter of synaptic vesicles	4584	99
1394	AF076249	Homo sapiens	zinc finger protein SBBI21	3208	99
1395	G02224	Homo sapiens	Human secreted protein, SEQ ID NO: 6305.	299	75
1396	AC004809	Arabidopsis thaliana	Similar to	130	34
1398	AF242519	Homo sapiens	zinc finger protein SBZF3	181	66
1399	AL133396	Homo sapiens	dJ1068H6.4 (prion protein like protein doppel)	962	100
1400	Y48611	Homo sapiens	Human breast tumour-associated protein 72.	817	99
1401	AC004472	Homo sapiens	P1.11659_5	280	54
1402	X91489	Saccharomyces cerevisiae	putative HMG box	164	27

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1403	Y79222	Homo sapiens	Human transferase TRNSFS-14.	2842	100
1404	X81058	Mus musculus	tex261	1010	99
1405	AB012084	Mus musculus	ITM	194	29
1406	AB030251	Homo sapiens	GTPase activating protein	3233	99
1407	AJ010585	Rattus rattus	PTB-like protein	2684	99
1408	X75760	Drosophila melanogaster	LRR47	364	29
1409	U76618	Mus musculus	N-RAP	804	48
1410	AC005578	Homo sapiens	P20887_1, partial CDS	835	63
1411	AE000284	Escherichia coli	orf, hypothetical protein.	360	100
1412	X01563	Escherichia coli	L5 (rplE) (aa 1-179)	911	100
1413	W78279	Homo sapiens	Fragment of human secreted protein encoded by gene 33.	1264	99
1414	AB031051	Homo sapiens	organic anion transporter OATP-E	3832	100
1415	M17466	Homo sapiens	coagulation factor XII	3455	100
1416	AF097994	Homo sapiens	L-kynurenine/alpha-aminoadipate aminotransferase	2202	99
1417	AF151077	Homo sapiens	HSPC243	1262	99
1418	Y09945	Rattus norvegicus	putative integral membrane transport protein	1098	61
1419	U13152	Mesocricetus auratus	guanine nucleotide-binding protein beta 5	2179	76
1420	AL162458	Homo sapiens	bA465L10.5 (KIAA1176 (novel protein, presumed ortholog of mouse K-Cl cotransporter KCC2))	5696	100
1421	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	152	29
1422	Y94923	Homo sapiens	Human secreted protein clone qs14_3 protein sequence SEQ ID NO:52.	4039	99
1423	AF177388	Homo sapiens	cancer-amplified transcriptional coactivator ASC-2	10748	99
1424	Y48517	Homo sapiens	Human breast tumour-associated protein 62.	1851	99
1425	AF208848	Homo sapiens	BM-006	1454	89
1426	AF208848	Homo sapiens	BM-006	853	79
1427	AF112886	Bos taurus	differentiation enhancing factor 1	4693	95
1428	U41387	Homo sapiens	Gu protein	1372	63
1429	AF161534	Homo sapiens	HSPC049	2853	78
1430	AF125043	Mus musculus	bisphosphate 3'-nucleotidase	275	30
1431	Y66718	Homo sapiens	Membrane-bound protein PRO1106.	1886	100
1432	AF193613	Homo sapiens	cell recognition molecule Caspr2	568	100
1433	AB044560	Mus musculus	Gliacolin	192	34
1434	R99900	Homo sapiens	NTII-1 nerve protein, facilitates regeneration of nerve cells.	707	51
1435	AF220530	Homo sapiens	myo-inositol 1-phosphate synthase A1	2904	100
1436	X70944	Homo sapiens	PTB-associated splicing factor	1261	72
1437	AF271732	Homo sapiens	bridging integrator-3	1282	100
1438	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	595	98
1439	AJ293659	Homo sapiens	mucolipidin	628	97
1440	AF219138	Homo sapiens	GGA3 long isoform	3083	100
1441	AF219138	Homo sapiens	GGA3 long isoform	3346	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1442	AB039669	Homo sapiens	ALEX3	1944	100
1443	AF237711	Drosophila melanogaster	Diablo	191	27
1444	AJ011896	Homo sapiens	Naf1 beta protein	439	39
1445	X73874	Homo sapiens	phosphorylase kinase	6233	98
1446	AF214124	Homo sapiens	breast carcinoma-associated antigen BCAA	3999	99
1447	AF003924	Homo sapiens	ANC_2H01	2645	99
1448	AF003136	Caenorhabditis elegans	contains weak similarity to an AMP-binding motif	2843	52
1449	AF155112	Homo sapiens	NY-REN-50 antigen	1184	89
1450	Y95004	Homo sapiens	Human secreted protein vc54_1, SEQ ID NO:48.	985	100
1451	AF107203	Homo sapiens	ataxin 2-binding protein	688	57
1452	AF107203	Homo sapiens	ataxin 2-binding protein	456	78
1453	Z38011	Mus musculus	DMR-N9	882	56
1454	X90568	Homo sapiens	Protein sequence and annotation available soon via LABEIT@EMBL-Heidelberg.DE	510	28
1455	AL035409	Homo sapiens	dJ564M11.3 (similar to sialyltranferase)	1356	100
1456	D44480	Mus musculus	MATH-2 protein	272	100
1458	AF141326	Homo sapiens	RNA helicase HDB/DICE1	478	45
1459	AF242552	Gallus gallus	retinovin	945	34
1460	U11036	Homo sapiens	Ibd1	724	84
1461	AB025258	Mus musculus	granuphilin-a	545	39
1462	Y08134	Homo sapiens	acid sphingomyelinase-like phosphodiesterase	2428	99
1463	AC004997	Homo sapiens	match to ESTs Z43979 (NID:g573097), R19699 (NID:g774333)	869	98
1464	AC004997	Homo sapiens	match to ESTs Z43979 (NID:g573097), R19699 (NID:g774333)	869	98
1465	U32743	Haemophilus influenzae Rd	fucose operon protein (fucU)	315	50
1466	Y09022	Homo sapiens	Not56-like protein	2342	100
1467	AC003034	Homo sapiens	Homolog of rat kidney-specific (KS) gene	1072	99
1468	AF071544	Spinacia oleracea	ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit N-methyltransferase I	333	26
1469	Y57930	Homo sapiens	Human transmembrane protein HTMPN-54.	1053	100
1470	AF032666	Rattus norvegicus	rsec5	4504	93
1471	Y70467	Homo sapiens	Human membrane channel protein-17 (MECHP-17).	452	74
1472	AL031033	Homo sapiens	C321D2.1 (Ribosomal Large Subunit Pseudouridine Synthase protein)	1694	100
1473	AF177292	Homo sapiens	genethonin 3	4026	98
1474	S45936	Homo sapiens	HTS1	1101	50
1475	Y86241	Homo sapiens	Human secreted protein HOABR60, SEQ ID NO:156.	1879	98
1476	AJ010317	Fugu rubripes	Sand	1278	68
1477	U42831	Caenorhabditis elegans	coded for by C. elegans cDNA yk99b4.3; similar to human transforming protein (PIR:S22157)	846	44
1478	X62447	Homo sapiens	PR 264	543	61
1479	X82209	Homo sapiens	MN1	7116	100
1480	U10536	Pan paniscus	MHC class I A	675	84

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1481	AL078599	Homo sapiens	dJ991C6.1 (novel protein similar to C. elegans F55A12.9 (Tr:P91086))	1274	65
1482	Z98977	Schizosaccharomyces pombe	putative vacuolar protein	256	29
1483	AB005662	Mus musculus	JNK/SAPK-associated protein-1	4968	92
1484	AL050120	Homo sapiens	hypothetical protein	716	100
1485	M27878	Homo sapiens	DNA binding protein	1006	53
1486	Y69161	Homo sapiens	Amino acid sequence of a partial protein kinase.	575	99
1487	X84156	Saccharomyces cerevisiae	ATH1	341	29
1488	AF038963	Homo sapiens	RNA helicase	446	34
1489	U56966	Caenorhabditis elegans	coded for by C. elegans cDNA yk30b3.5; coded for by C. elegans cDNA yk30b3.3	620	42
1490	AE000989	Archaeoglobus fulgidus	enoyl-CoA hydratase (fad-4)	533	46
1491	M80633	Rattus norvegicus	adenylyl cyclase type IV	707	95
1492	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	3513	99
1493	Y17220	Homo sapiens	Human secreted protein (clone fj283-11).	462	37
1494	AF133670	Mus musculus	ARL-6 interacting protein-2	701	97
1495	Y94897	Homo sapiens	Human protein clone HP10574.	1371	100
1496	AL049699	Homo sapiens	dJ747H23.2 (novel protein)	1550	100
1497	AF037447	Homo sapiens	ribosomal S6 protein kinase	2427	100
1498	AL445067	Thermoplasma acidophilum	putative target YPL207w of the HAP2 transcriptional complex related protein	269	35
1499	AB039947	Homo sapiens	X11L-binding protein 51	227	36
1500	AJ277750	Homo sapiens	UBASH3A protein	3509	100
1501	AL050333	Homo sapiens	dJ93K22.1 (novel protein (contains DKFZP564B116))	2439	100
1502	AF179896	Homo sapiens	TALE homeobox protein Meis2b	1140	100
1503	AF178948	Homo sapiens	TALE homeobox protein Meis2a	1177	100
1504	Y53005	Homo sapiens	Human secreted protein clone pn749_8 protein sequence SEQ ID NO:16.	1442	99
1505	X82494	Homo sapiens	fibulin-2	3580	99
1506	X98296	Homo sapiens	ubiquitin hydrolase	783	42
1507	AL034548	Homo sapiens	dJ1103G7.6 (novel protein)	1098	100
1508	Y76144	Homo sapiens	Human secreted protein encoded by gene 21.	1736	100
1509	AF220182	Homo sapiens	uncharacterized hypothalamus protein HT008	1181	98
1510	U64601	Caenorhabditis elegans	Gene probably begins in the next cosmid	415	58
1511	AL356192	Neurospora crassa	related to MDM1 protein	196	29
1512	D17629	Homo sapiens	N-acetylgalactosamine 6-sulfate sulfatase (GALNS)	1829	100
1513	AF168717	Homo sapiens	x 009 protein	694	99
1514	AJ243531	Homo sapiens	nM15 protein	735	100
1515	AC003672	Arabidopsis thaliana	putative C3HC4-type RING zinc finger protein	407	30
1516	AF115435	Rattus norvegicus	syntaxin 17	1374	90
1517	AF003140	Caenorhabditis elegans	C44E4.5 gene product	274	31
1518	AB002584	Rattus norvegicus	beta-alanine-pyruvate aminotransferase	2238	82
1519	AL121764	Schizosaccha	yeast atp12 protein precursor	270	30

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		romyces pombe	homolog		
1520	AF255910	Homo sapiens	vascular endothelial junction-associated molecule	547	100
1521	D31764	Homo sapiens	KIAA0064	170	27
1522	Y66634	Homo sapiens	Membrane-bound protein PRO190.	985	100
1523	Y94450	Homo sapiens	Human inflammation associated protein	250	43
1524	AC000107	Arabidopsis thaliana	F17F8.22	277	37
1525	AF109377	Mus musculus	ldlBp	1277	83
1526	AL031427	Homo sapiens	dJ167A19.4 (novel protein)	1432	99
1527	Y08135	Mus musculus	acid sphingomyelinase-like phosphodiesterase	1496	79
1528	AK024423	Homo sapiens	FLJ00012 protein	611	100
1529	AF154502	Homo sapiens	quiescent cell proline dipeptidase	679	100
1530	AF205598	Homo sapiens	transposase-like protein	1368	100
1531	AF251039	Homo sapiens	putative zinc finger protein	1420	50
1532	W74805	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24.	493	57
1533	AF039023	Homo sapiens	Ran-GTP binding protein; RanBP6	5707	99
1534	AC007190	Arabidopsis thaliana	F23N19.9	374	37
1535	AB027564	Homo sapiens	DINB1	4482	100
1536	Y36178	Homo sapiens	Human secreted protein	377	87
1537	Y50907	Homo sapiens	Human fetal brain cDNA clone vb3_1 derived protein.	3693	99
1538	AF017368	Mus musculus	faciogenital dysplasia protein 2	177	47
1539	AF266756	Homo sapiens	sphingosine kinase	2011	99
1540	Z48804	Homo sapiens	OA1	2238	100
1541	AF000195	Caenorhabditis elegans	Contains similarity to Pfam domain: PF00169 (PH), Score=20.6, E-value=1.9e-05, N=1	379	42
1542	Y71159	Homo sapiens	Human phosphodiesterase interacting protein, myomegalin.	9415	99
1543	X76092	Homo sapiens	DNA binding protein RFX3	3327	100
1544	AB015330	Homo sapiens	HRIHFB2007	631	50
1545	AF198487	Homo sapiens	transcription factor LBP-1b	2822	100
1546	AF016417	Caenorhabditis elegans	Similar to BZIP transcription factor	518	42
1547	X55885	Homo sapiens	KDEL receptor	1106	100
1548	AB035495	Carassius auratus	ubiquitin-activating enzyme E1	836	42
1549	AL021707	Homo sapiens	dJ508I15.4 (KIAA0668)	3688	100
1550	AJ223978	Bacillus subtilis	YvqK protein	292	42
1551	AF145615	Drosophila melanogaster	BcDNA.GH03377	822	44
1552	AL157734	Schizosaccharomyces pombe	putative mannosyltransferase involved in N-glycosylation	435	37
1553	AF079527	Mus musculus	IER5	691	63
1554	AB026291	Rattus norvegicus	acetoacetyl-CoA synthetase	1099	88
1555	Y44722	Homo sapiens	Human immune system molecule, ISMO-3.	1780	99
1556	AF116553	Drosophila melanogaster	antennal-specific short-chain dehydrogenase/reductase	277	32
1557	Y71056	Homo sapiens	Human membrane transport	1975	99

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			protein, MTRP-1.		
1558	Y71056	Homo sapiens	Human membrane transport protein, MTRP-1.	1975	99
1559	Y71056	Homo sapiens	Human membrane transport protein, MTRP-1.	1894	97
1560	AF092050	Mus musculus	beta-1,3-N-acetylglucosaminyltransferase	262	44
1561	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4)))	1607	97
1562	AJ131890	Homo sapiens	DNA polymerase lambda	3002	100
1563	AL035424	Homo sapiens	dA22D12.1 (novel protein similar to Drosophila Kelch proteins)	3015	100
1564	AC002400	Homo sapiens	Gene product with similarity to Ubiquitin binding enzyme	2790	100
1565	AC005306	Homo sapiens	R27216_1	919	82
1566	AF000195	Caenorhabditis elegans	Contains similarity to Pfam domain: PF00169 (PH), Score=20.6, E-value=1.9e-05, N=1	550	45
1567	AB033281	Homo sapiens	F-box and WD-repeats protein beta-TRCP2 isoform C	2879	100
1568	D49473	Mus musculus	truncated form of Sox17	1047	78
1569	AK025270	Homo sapiens	unnamed protein product	210	91
1570	X75756	Homo sapiens	protein kinase C mu	4797	99
1571	AF145713	Homo sapiens	SCHIP-1	2388	100
1572	AE003831	Drosophila melanogaster	CG18445 gene product	180	31
1573	AF074603	Streptomyces griseus subsp. griseus	NonF	205	38
1574	U28993	Caenorhabditis elegans	F22D3.3 gene product	144	27
1575	AF129507	Homo sapiens	transcription factor ICBP90	287	68
1576	X64878	Homo sapiens	oxytocin receptor	2002	100
1577	AF237711	Drosophila melanogaster	Diablo	421	54
1578	G00975	Homo sapiens	Human secreted protein, SEQ ID NO: 5056.	480	100
1579	AF248744	Cryptosporidium parvum	thrombospondin-related adhesive protein	123	33
1580	AL121782	Homo sapiens	dJ585I14.2 (novel protein (translation of cDNA Em:AK000219))	663	100
1581	AF041853	Homo sapiens	kinesin family member protein KIF3A	345	33
1582	AF025441	Homo sapiens	Opa-interacting protein OIP5	1198	100
1583	AE001803	Thermotoga maritima	glycerate kinase, putative	349	34
1584	AF252283	Homo sapiens	Kelch-like 1 protein	3973	100
1585	AF169675	Homo sapiens	leucine-rich repeat transmembrane protein FLRT1	3494	99
1586	AF118274	Homo sapiens	DNb-5	2628	97
1587	X79440	Homo sapiens	NADP+-dependent malic enzyme	3167	99
1588	X99802	Homo sapiens	ZYG homologue	3966	99
1589	AF169803	Homo sapiens	flavohemoprotein b5+b5R	2563	100
1590	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	181	47
1591	Z25535	Homo sapiens	nuclear pore complex protein hnup153	7567	99
1592	X13293	Homo sapiens	B-myb protein (AA 1-700)	3678	99
1593	M74027	Homo sapiens	mucin	242	27
1594	AL139314	Schizosaccharomyces	hypothetical protein	235	54

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		pombe			
1595	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1596	Y94906	Homo sapiens	Human secreted protein clone rb649_3 protein sequence SEQ ID NO:18.	2236	98
1597	AF174605	Homo sapiens	F-box protein Fbx25	1408	99
1598	AB032254	Homo sapiens	bromodomain adjacent to zinc finger domain 2A	9676	98
1599	X73114	Homo sapiens	slow MyBP-C	5568	95
1600	X82200	Homo sapiens	gpStaf50	2305	100
1601	Y00876	Homo sapiens	Human LAPH-1 protein sequence.	1149	98
1602	AJ223351	Homo sapiens	HIRA-interacting protein 3	2821	99
1603	AJ222801	Homo sapiens	neutral sphingomyelinase	2268	99
1604	AJ222801	Homo sapiens	neutral sphingomyelinase	1601	99
1605	AF185576	Mus musculus	POZ/zinc finger transcription factor ODA-8	3435	97
1606	AF093744	Homo sapiens	unknown	131	100
1607	A12142	synthetic construct	IFN-pseudo-omega 2	800	98
1608	Y57949	Homo sapiens	Human transmembrane protein HTMPN-73.	1868	100
1609	AF151044	Homo sapiens	HSPC210	681	97
1610	X15218	Homo sapiens	ski protein (AA 1 - 728)	3765	100
1611	Y08200	Homo sapiens	rab geranylgeranyl transferase	2976	100
1612	AF220560	Homo sapiens	B/K protein	2486	99
1613	AC004481	Arabidopsis thaliana	nodulin-like protein	371	26
1614	Y09501	Homo sapiens	NADH-cytochrome-b5 reductase	1607	100
1615	Y15521	Homo sapiens	start position 1	3150	97
1616	AJ010750	Rattus norvegicus	Castration induced prostatic apoptosis related protein-1, (CIPAR-1)	890	62
1617	X58079	Homo sapiens	S100 alpha protein	481	100
1618	Y66678	Homo sapiens	Membrane-bound protein PRO1009.	967	100
1619	AJ242973	Homo sapiens	peptide methionine sulfoxide reductase	929	100
1620	AF150733	Homo sapiens	AD-014 protein	288	100
1621	AJ007509	Homo sapiens	E1B-55kDa-associated protein	4646	98
1622	X64177	Homo sapiens	metallothionein	380	100
1623	AE001045	Archaeoglobus fulgidus	A. fulgidus predicted coding region AF0859	240	36
1624	AL355013	Schizosaccharomyces pombe	mitochondrial carrier protein	403	34
1625	Y66746	Homo sapiens	Membrane-bound protein PRO1198.	1184	100
1626	D90053	Sus scrofa	destrin	863	100
1627	Y35954	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 203.	756	100
1628	AL031775	Homo sapiens	dJ30M3.2 (novel protein)	470	100
1629	AF132484	Mus musculus	unknown	286	68
1630	AF017096	Drosophila melanogaster	similar to C. elegans R10H10.6 and S. cerevisiae YD8419.03c	493	61
1631	X03077	Homo sapiens	lactate dehydrogenase-A	1704	100
1632	AF151084	Homo sapiens	HSPC250	763	100
1633	AJ001874	Homo sapiens	orf	255	97
1634	AC012187	Arabidopsis thaliana	Contains weak similarity to GATA-6 DNA-binding protein gb|H36135, gb|Z26200 come from this gene.	143	38

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1635	AF026246	Homo sapiens	HERV-E integrase	411	90
1636	Y50943	Homo sapiens	Human adult brain cDNA clone ve8_1 derived protein.	1126	95
1637	AF134593	Homo sapiens	L-pipecolic acid oxidase	2068	99
1638	AJ238247	Mus musculus	putative phosphatase subunit	1948	96
1639	Y94942	Homo sapiens	Human secreted protein clone yk251_1 protein sequence SEQ ID NO:90.	1320	100
1640	AF235030	Homo sapiens	BM88 antigen	766	99
1641	AF233288	Drosophila melanogaster	WDS	358	26
1642	M19351	Mus musculus	immunoglobulin heavy chain binding protein	145	34
1643	Y70452	Homo sapiens	Human membrane channel protein-2 (MECHP-2).	1352	100
1644	AF176520	Mus musculus	WD repeat-containing F-box protein FBW5	2676	88
1645	W67816	Homo sapiens	Human secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1646	X67155	Homo sapiens	mitotic kinase-like protein-1	4456	99
1647	M63180	Homo sapiens	threonyl-tRNA synthetase	1040	61
1648	Y87342	Homo sapiens	Human signal peptide containing protein HSPP-119 SEQ ID NO:119.	1566	93
1649	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	4137	100
1650	AC007136	Homo sapiens	Putative map kinase interacting kinase	856	99
1651	AB015346	Homo sapiens	Eps15R	4464	99
1652	AL161576	Arabidopsis thaliana	putative protein	1341	48
1653	AC005313	Arabidopsis thaliana	putative calmodulin	288	28
1654	AL031428	Homo sapiens	dJ184J9.1 (KIAA0601 protein)	3526	100
1655	AL031428	Homo sapiens	dJ184J9.1 (KIAA0601 protein)	3526	100
1656	AB017910	Dictyostelium discoideum	myoM	297	32
1657	Y28919	Homo sapiens	Human regulatory protein HRGP-5.	2251	99
1658	AF056191	Homo sapiens	TPA inducible protein	2744	98
1659	U76846	Arabidopsis thaliana	ubiquitin-specific protease	137	35
1660	AL078627	Schizosaccharomyces pombe	actin-like protein; (2 actin domains)	320	34
1662	X52022	Homo sapiens	collagen type VI, alpha 3 chain	16274	99
1663	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit 4	1811	100
1664	AF214736	Homo sapiens	EH domain containing protein 2	2774	100
1665	Z48613	Saccharomyces cerevisiae	unknown	138	26
1666	AF177385	Homo sapiens	cytochrome c oxidase assembly protein isoform 2	1395	99
1667	AC007842	Homo sapiens	BC331191_1	1581	47
1668	S67513	Borna disease virus BDV, WT-1, Halle B1/91, horse brain, field isolate, Peptide, 370	p40	397	43

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		aa			
1669	Z99753	Schizosaccharomyces pombe	putative NOL1-NOP2-sun family nucleolar protein	569	47
1670	G03130	Homo sapiens	Human secreted protein, SEQ ID NO: 7211.	427	97
1671	M96625	Gallus gallus	cardiac muscle tensin	1185	54
1672	AF174482	Homo sapiens	polycomb 3	2005	99
1673	Y51846	Homo sapiens	Human 18.1 homolog protein fragment.	233	29
1674	AF255334	Homo sapiens	EXP35	152	29
1675	Y94867	Homo sapiens	Human protein clone HP10563.	109	30
1676	Y25712	Homo sapiens	Human secreted protein encoded from gene 2.	3043	99
1677	Y25712	Homo sapiens	Human secreted protein encoded from gene 2.	1580	91
1678	AF163151	Homo sapiens	dentin sialophosphoprotein precursor	170	17
1679	AF163151	Homo sapiens	dentin sialophosphoprotein precursor	170	17
1680	AK024453	Homo sapiens	FLJ00045 protein	1349	100
1681	AF019236	Dictyostelium discoideum	TipD	613	34
1682	AJ243459	Leishmania major	proteophosphoglycan	153	26
1683	Z69369	Schizosaccharomyces pombe	putative GTP-binding protein	560	46
1684	X94910	Homo sapiens	ERp28	1334	100
1685	AF286475	Takifugu rubripes	retinitis pigmentosa GTPase regulator-like protein	196	19
1686	AF191298	Homo sapiens	vacuolar sorting protein 35	4087	100
1687	AJ275986	Homo sapiens	transcription factor	2958	100
1688	AJ275986	Homo sapiens	transcription factor	1886	88
1689	X07311	Drosophila melanogaster	heat shock protein	138	43
1690	AF240463	Rattus norvegicus	LIS1-interacting protein NUDE1	1383	83
1691	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	1256	68
1692	AJ272079	Homo sapiens	APOBEC-1 stimulating protein	1336	60
1693	AF177942	Xenopus laevis	katanin p60	1664	66
1694	AF263539	Homo sapiens	arginine N-methyltransferase	1774	100
1695	AF222689	Homo sapiens	protein arginine N-methyltransferase 1-variant 2	1182	81
1696	AK000193	Homo sapiens	unnamed protein product	1060	100
1697	AB041035	Homo sapiens	kidney superoxide-producing NADPH oxidase	3122	100
1698	AB041035	Homo sapiens	kidney superoxide-producing NADPH oxidase	2181	100
1699	AF025772	Homo sapiens	C2H2 zinc finger protein	488	54
1700	Y44676	Homo sapiens	Human ARF-Related Protein-1 (HARP-1).	938	97
1701	AK022407	Homo sapiens	unnamed protein product	315	98
1702	AB024574	Homo sapiens	GTP-binding like protein 2	1172	100
1703	AF055078	Homo sapiens	zinc finger protein 42	421	52
1704	AF198092	Mus musculus	RP42	1057	77
1705	AE003573	Drosophila melanogaster	CG12474 gene product	161	33
1706	AB036345	Drosophila melanogaster	aquaporin	164	24
1707	Y55927	Homo sapiens	Human STLK2 protein.	2146	100
1708	U27121	Danio rerio	G12	212	47
1709	AL391710	Arabidopsis	putative protein	505	50

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		thaliana			
1710	B01311	Homo sapiens	Human PRO241 polypeptide.	1649	97
1711	U40750	Mus musculus	formin binding protein 30	4561	85
1712	AJ011118	Mus musculus	skeletal muscle and cardiac protein	1490	89
1713	AF255303	Homo sapiens	membrane-associated nucleic acid binding protein	4416	99
1714	AF255303	Homo sapiens	membrane-associated nucleic acid binding protein	2960	100
1715	U08227	Rattus norvegicus	Ras-related protein	511	51
1716	AF168795	Rattus norvegicus	schlafen-4	1129	44
1717	AF196304	Homo sapiens	SUMO-1-specific protease	5804	99
1718	AL355737	Homo sapiens	HMG20A	1782	100
1719	AB029333	Halocynthia roretzi	HrPET-1	1069	46
1720	AF071317	Mus musculus	COP9 complex subunit 7b	1297	97
1721	AJ272215	Homo sapiens	HEYL protein	1681	99
1722	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
1723	AL032643	Caenorhabditis elegans	similar to Uncharacterized protein family UPF0034,	825	41
1724	G01972	Homo sapiens	Human secreted protein, SEQ ID NO: 6053.	586	92
1725	Y94441	Homo sapiens	Human Adipose Specific Protein 1.	1231	100
1726	AF255443	Homo sapiens	CGI-201 protein	4397	99
1727	AF183426	Homo sapiens	HT004 protein	1810	99
1728	D10884	Bos taurus	neurocalcin	1002	99
1729	Z18529	Gallus gallus	tensin	1411	84
1730	Z73423	Caenorhabditis elegans	cDNA EST EMBL:Z14908 comes from this gene-cDNA EST this gene	233	41
1732	AF090891	Homo sapiens	PRO0105	470	30
1733	AJ277724	Homo sapiens	histone deacetylase 8	2015	100
1734	G04050	Homo sapiens	Human secreted protein, SEQ ID NO: 8131.	503	95
1735	D45913	Mus musculus	leucine-rich-repeat protein	3531	94
1736	AF096709	Drosophila virilis	failed axon connections protein	276	32
1737	AF195120	Homo sapiens	dynactin p62 subunit	2417	99
1738	L15314	Caenorhabditis elegans	contains similarity to Pfam family PF01772 N=1	206	37
1739	X54618	Listeria monocytogenes	phosphadidylinositol specific phospholipase C	134	27
1740	AL031658	Homo sapiens	dJ310O13.4 (novel protein similar to predicted C. elegans an C. intestinalis proteins)	123	31
1741	Y35924	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 173.	1013	99
1742	AC013354	Arabidopsis thaliana	F15H18.15	202	32
1743	W75771	Homo sapiens	Human GTP binding protein APD08.	1932	59
1744	W75771	Homo sapiens	Human GTP binding protein APD08.	1854	61
1745	AF221098	Homo sapiens	Ral guanine nucleotide exchange factor RalGPS1A	1224	70
1746	Y99372	Homo sapiens	Human PRO1430 (UNQ736) amino acid sequence SEQ ID NO:116.	1332	99
1747	Y94294	Homo sapiens	Human coenzyme A-utilising	842	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			enzyme CoARN-2.		
1748	AK024436	Homo sapiens	FLJ00026 protein	1619	100
1749	AE000877	Methanobacterium thermoautotrophicum	conserved protein	231	36
1750	AF101361	Drosophila melanogaster	Abnormal X segregation	193	33
1751	Y15067	Homo sapiens	ZNF232	889	100
1752	AF251038	Homo sapiens	GAP-like protein	822	100
1753	AC003093	Homo sapiens	OXYSTEROL-BINDING PROTEIN; 45% similarity to P22059 (PID:g129308)	352	57
1754	X69089	Homo sapiens	165kD protein	5703	99
1755	AL049795	Homo sapiens	dJ622L5.3 (novel protein)	1039	100
1756	AL031393	Homo sapiens	dJ733D15.1 (Zinc-finger protein)	2765	100
1757	AB040672	Homo sapiens	UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase	2020	99
1758	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	776	43
1759	AF117653	Homo sapiens	double homeobox protein	375	54
1760	Y12065	Homo sapiens	hNop56	2959	99
1761	AL049712	Homo sapiens	dJ686C3.2 (nucleolar protein hNop56)	2595	99
1762	AC002394	Homo sapiens	Gene product with similarity to dynein beta subunit	1542	51
1763	AF169017	Homo sapiens	formiminotransferase cyclodeaminase	877	100
1764	U91541	Homo sapiens	human formiminotransferase cyclodeaminase (ftcd)protein, carboxy-terminal end	596	100
1765	AB013365	Bacillus halodurans	YlqF	350	34
1766	Y38421	Homo sapiens	Human secreted protein encoded by gene No. 36.	145	71
1767	AC009176	Arabidopsis thaliana	putative ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit N-methyltransferase I	216	27
1768	AK000647	Homo sapiens	unnamed protein product	737	99
1769	AJ238982	Homo sapiens	VNN3 protein	2665	99
1770	U73522	Homo sapiens	AMSH	1214	56
1771	U89435	Mus musculus	unknown	829	86
1772	S70011	Rattus sp.	tricarboxylate carrier	1604	95
1773	AL035086	Homo sapiens	dJ44A20.2 (novel protein)	2036	100
1774	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	1057	99
1775	AF110330	Homo sapiens	glutaminase	3146	100
1776	AJ269529	Homo sapiens	glycerol 3-phosphate permease	2787	100
1777	Z81579	Caenorhabditis elegans	cDNA EST yk76f1.5 comes from this gene	232	31
1778	AY007239	Homo sapiens	monooxygenase X	1875	99
1779	AL109608	Schizosaccharomyces pombe	oxysterol-binding protein family	644	38
1780	AF254260	Homo sapiens	tuftelin 1	1729	100
1781	L07924	Mus musculus	guanine nucleotide dissociation stimulator	247	50
1782	AF295773	Homo sapiens	ral guanine nucleotide dissociation stimulator	142	49
1783	AK024475	Homo sapiens	FLJ00068 protein	4333	100
1784	AK024475	Homo sapiens	FLJ00068 protein	3996	93
1785	G03933	Homo sapiens	Human secreted protein, SEQ ID NO: 8014.	570	100
1786	S82637	Homo sapiens	Ig lambda-like gene/beta-	247	100

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			glucuronidase exon 11 homolog		

TRADOCS:1416280.1(%CT401!.DOC)

TABLE 3

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
2	BL00240	Receptor tyrosine kinase class III proteins.	BL00240B 24.70 8.250e-12 157-181
3	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109D 17.04 8.085e-13 358-381
4	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 9.400e-10 1129-1146 BL00028 16.07 1.257e-09 820-837
5	BL00023	Type II fibronectin collagen-binding domain proteins.	BL00023 24.31 8.920e-33 413-450 BL00023 24.31 4.545e-27 353-390
6	BL00023	Type II fibronectin collagen-binding domain proteins.	BL00023 24.31 8.920e-33 413-450 BL00023 24.31 4.545e-27 353-390
7	BL00023	Type II fibronectin collagen-binding domain proteins.	BL00023 24.31 8.920e-33 413-450 BL00023 24.31 4.545e-27 353-390
8	BL00023	Type II fibronectin collagen-binding domain proteins.	BL00023 24.31 8.920e-33 413-450 BL00023 24.31 4.545e-27 353-390
9	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 5.119e-09 863-917
10	PR00464	E-CLASS P450 GROUP II SIGNATURE	PR00464D 17.40 6.182e-12 294-312 PR00464G 12.41 4.231e-11 377-393
11	PR00734	GLYCOSYL HYDROLASE FAMILY 7 SIGNATURE	PR00734I 11.46 4.296e-09 502-520
12	PF00023	Ank repeat proteins.	PF00023B 14.20 6.500e-10 89-99 PF00023B 14.20 2.636e-09 56-66
14	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 3.848e-09 79-113
15	PR00208	GLIADIN AND LMW GLUTENIN SUPERFAMILY SIGNATURE	PR00208A 12.59 9.868e-10 517-535 PR00208A 12.59 2.233e-09 520-538
17	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 8.200e-14 282-295 PD00066 13.92 9.400e-14 477-490 PD00066 13.92 6.500e-13 505-518 PD00066 13.92 9.500e-13 254-267 PD00066 13.92 1.429e-12 393-406 PD00066 13.92 6.571e-12 421-434
18	BL00845	CAP-Gly domain proteins.	BL00845 16.43 2.200e-25 55-80
20	BL00487	IMP dehydrogenase / GMP reductase proteins.	BL00487E 16.12 5.737e-26 154-199 BL00487F 18.79 8.984e-22 235-276 BL00487G 26.82 4.082e-12 287-329
21	BL00487	IMP dehydrogenase / GMP reductase proteins.	BL00487E 16.12 5.737e-26 154-199 BL00487F 18.79 8.984e-22 235-276 BL00487G 26.82 4.082e-12 348-390
22	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 3.250e-26 302-333

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
23	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 3.250e-26 302-333
25	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115T 8.45 7.273e-29 1208-1242 BL00115Q 18.08 2.776e-21 953-983 BL00115Y 11.86 8.000e-17 1604-1650 BL00115M 19.19 8.130e-16 731-774 BL00115H 14.34 9.392e-16 463-496 BL00115A 15.44 7.414e-15 43-82 BL00115R 6.50 6.128e-14 983-1010 BL00115J 16.71 9.289e-14 591-617 BL00115I 8.33 4.336e-13 535-590 BL00115L 12.25 5.939e-13 662-694 BL00115G 11.65 6.011e-13 435-463 BL00115K 15.03 3.417e-10 617-659 BL00115O 16.76 5.805e-10 863-913 BL00115P 11.54 7.538e-10 913-953 BL00115S 18.24 7.968e-10 1010-1052 BL00115U 10.34 4.475e-09 1242-1265
26	BL00420	Speract receptor repeat proteins domain proteins.	BL00420A 20.42 4.109e-11 81-110 BL00420A 20.42 8.820e-10 84-113
27	BL00050	Ribosomal protein L23 proteins.	BL00050A 23.71 9.250e-27 94-127 BL00050B 14.81 8.125e-12 133-147
28	PR00925	NONHISTONE CHROMOSOMAL PROTEIN HMG17 FAMILY SIGNATURE	PR00925B 3.73 3.089e-10 41-54
29	PF00756	Putative esterase.	PF00756C 14.12 1.108e-09 486-516
32	BL00557	FMN-dependent alpha-hydroxy acid dehydrogenases proteins.	BL00557D 17.76 5.065e-37 274-316 BL00557A 35.08 8.909e-29 24-73 BL00557C 15.59 1.000e-28 227-257 BL00557B 21.27 8.898e-22 130-169
34	PR00629	SHC PHOSPHOTYROSINE INTERACTION DOMAIN SIGNATURE	PR00629E 9.90 5.886e-35 299-328 PR00629F 10.95 8.364e-32 334-361 PR00629B 13.66 3.786e-27 224-247 PR00629A 13.45 8.364e-21 206-222 PR00629C 3.80 4.000e-12 249-261 PR00629D 12.45 3.739e-11 276-286
35	PD01270	RECEPTOR FC IMMUNOGLOBULIN AFFIN.	PD01270A 17.22 1.000e-40 39-79 PD01270B 22.18 2.875e-38 94-131 PD01270D 24.66 3.700e-34 171-207 PD01270C 19.54 3.455e-30 137-166
36	PD01270	RECEPTOR FC IMMUNOGLOBULIN AFFIN.	PD01270A 17.22 1.000e-40 39-79 PD01270B 22.18 2.875e-38 94-131

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			PD01270D 24.66 3.700e-34 171-207 PD01270C 19.54 3.455e-30 137-166
37	BL00412	Neuromodulin (GAP-43) proteins.	BL00412C 10.28 9.241e-10 264-298
38	BL00412	Neuromodulin (GAP-43) proteins.	BL00412C 10.28 9.241e-10 264-298
39	BL00412	Neuromodulin (GAP-43) proteins.	BL00412C 10.28 9.241e-10 264-298
40	PR00380	KINESIN HEAVY CHAIN SIGNATURE	PR00380B 12.64 7.366e-14 342-360 PR00380C 13.18 6.927e-13 375-394 PR00380D 9.93 2.180e-12 429-451 PR00380A 14.18 5.154e-12 143-165
44	BL00345	Ets-domain proteins.	BL00345B 21.28 1.000e-40 239-290 BL00345A 13.96 2.452e-14 204-223
45	BL00345	Ets-domain proteins.	BL00345B 21.28 1.000e-40 215-266 BL00345A 13.96 2.452e-14 180-199
46	DM01551	kw OSTEOINDUCTIVE YOPM MEMBRANE OUTER.	DM01551A 15.63 3.538e-26 172-202 DM01551C 14.62 3.571e-17 232-252 DM01551B 8.84 4.750e-11 214-226
47	PR00876	NEMATODE METALLOTHIONEIN SIGNATURE	PR00876B 7.66 9.328e-11 246-260
48	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 4.231e-33 6-45
50	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972D 22.55 7.750e-19 994-1019 BL00972A 11.93 7.120e-18 216-234 BL00972E 20.72 9.471e-14 1020-1042 BL00972C 16.48 7.000e-13 360-375 BL00972B 9.45 8.269e-10 302-312
51	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972D 22.55 7.750e-19 990-1015 BL00972A 11.93 7.120e-18 216-234 BL00972E 20.72 9.471e-14 1016-1038 BL00972C 16.48 7.000e-13 360-375 BL00972B 9.45 8.269e-10 302-312
52	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 3.063e-14 10-54
53	PR00988	URIDINE KINASE SIGNATURE	PR00988A 6.39 8.500e-17 20-38 PR00988F 12.23 7.828e-15 196-210 PR00988C 13.64 6.108e-14 104-120 PR00988E 8.27 3.872e-11 174-186 PR00988D 5.95 6.878e-10 160-171 PR00988B 11.60 2.915e-09 57-69
55	PR00762	CHLORIDE CHANNEL SIGNATURE	PR00762C 9.29 4.682e-21 294-314 PR00762D 11.29 4.103e-19 509-530 PR00762A 14.22 9.333e-18 199-217

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			PR00762F 15.12 3.100e-16 563-583 PR00762B 12.12 6.063e-16 230-250 PR00762E 12.07 2.286e-15 545-562 PR00762G 14.13 6.276e-13 601-616
56	BL00216	Sugar transport proteins.	BL00216B 27.64 8.800e-10 153-203
58	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 2.049e-10 1080-1135
59	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 2.049e-10 1062-1117
61	PD01929	KINASE TYPE RESISTANCE ANTIBIOTIC TRANSFERASE AM.	PD01929E 10.76 9.018e-09 206-221
68	PR00360	C2 DOMAIN SIGNATURE	PR00360A 14.59 7.395e-09 680-693
69	PR00360	C2 DOMAIN SIGNATURE	PR00360A 14.59 7.395e-09 670-683
70	PF00651	BTB (also known as BR-C/Ttk) domain proteins.	PF00651 15.00 8.714e-10 51-64
72	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 5.304e-09 108-118
73	BL00239	Receptor tyrosine kinase class II proteins.	BL00239B 25.15 7.075e-12 118-166
74	BL00790	Receptor tyrosine kinase class V proteins.	BL00790N 13.25 6.116e-10 93-120
76	DM00471	0 PROKARYOTIC DNA TOPOISOMERASE I.	DM00471A 11.73 9.357e-13 53-66 DM00471B 8.45 4.857e-12 70-81
80	PD02876	DECARBOXYLASE PHOSPHATIDYLSERINE.	PD02876C 8.80 2.723e-13 223-236 PD02876D 12.13 2.588e-12 334-351
81	PD02876	DECARBOXYLASE PHOSPHATIDYLSERINE.	PD02876C 8.80 2.723e-13 282-295 PD02876D 12.13 2.588e-12 393-410
83	BL00708	Prolyl endopeptidase family serine proteins.	BL00708B 24.91 7.197e-12 570-601
84	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014C 15.44 8.043e-09 985-1004
86	PR00678	PI3 KINASE P85 REGULATORY SUBUNIT SIGNATURE	PR00678H 9.13 1.379e-09 246-269
89	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320C 13.01 8.200e-09 264-279 PR00320B 12.19 8.650e-09 264-279
93	BL00455	Putative AMP-binding domain proteins.	BL00455 13.31 2.588e-14 316-332
95	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 4.000e-10 123-154
96	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 4.000e-10 212-243
97	PR00081	GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE	PR00081B 10.38 6.318e-13 134-146 PR00081A 10.53 2.500e-12 54-72
98	PR00380	KINESIN HEAVY CHAIN SIGNATURE	PR00380A 14.18 5.500e-24 401-423 PR00380D 9.93 7.188e-20 613-635 PR00380B 12.64 7.517e-16 529-547 PR00380C 13.18 2.756e-13 560-579

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
102	PR00300	ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT SIGNATURE	PR00300A 9.56 7.545e-14 289-308
104	BL00479	Phorbol esters / diacylglycerol binding domain proteins.	BL00479B 12.57 6.786e-18 298-314 BL00479A 19.86 4.913e-16 155-178 BL00479A 19.86 4.300e-13 272-295 BL00479B 12.57 6.294e-12 181-197
106	BL01019	ADP-ribosylation factors family proteins.	BL01019A 13.20 8.013e-12 43-83
107	DM01970	0 kw ZK632.12 YDR313C ENDOSOMAL III.	DM01970B 8.60 5.000e-16 403-416
108	BL00191	Cytochrome b5 family, heme-binding domain proteins.	BL00191K 17.38 4.951e-27 238-282 BL00191J 11.37 6.447e-17 182-204
109	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 4.938e-37 8-47
110	BL01138	Scorpion short toxins proteins.	BL01138A 10.96 8.297e-10 38-50
113	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 5.800e-23 156-187 BL00107B 13.31 9.100e-14 225-241
117	BL00214	Cytosolic fatty-acid binding proteins.	BL00214B 26.51 1.000e-17 46-91 BL00214A 21.17 7.052e-11 5-31
118	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 8.560e-13 36-67
119	PR00529	GONADOTROPHIN RELEASING HORMONE RECEPTOR SIGNATURE	PR00529C 11.03 7.506e-10 158-177
120	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320C 13.01 9.400e-09 80-95
121	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320C 13.01 9.400e-09 80-95
127	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 7.158e-13 216-241
128	BL01032	Protein phosphatase 2C proteins.	BL01032C 6.14 3.195e-12 147-157 BL01032H 11.25 5.680e-11 318-331 BL01032G 8.33 8.932e-11 282-296 BL01032I 10.42 8.902e-09 379-389
129	BL01310	ATP1G1 / PLM / MAT8 family proteins.	BL01310 14.74 6.694e-26 28-64
130	PR00990	RIBOKINASE SIGNATURE	PR00990B 12.32 9.534e-15 47-67 PR00990A 16.23 5.500e-14 20-42 PR00990C 12.62 2.412e-09 119-133
133	BL00880	Acyl-CoA-binding protein.	BL00880 17.52 5.575e-26 72-122
134	BL00030	Eukaryotic RNA-binding region RNP-1 proteins.	BL00030A 14.39 9.308e-14 18-37
135	PR00215	NEUROMODULIN SIGNATURE	PR00215C 13.98 6.779e-10 475-496
136	BL01310	ATP1G1 / PLM / MAT8 family proteins.	BL01310 14.74 2.432e-29 71-107
140	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 7.882e-14 214-231 BL00028 16.07 9.471e-14 102-119 BL00028 16.07 2.800e-13 18-35

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			BL00028 16.07 5.500e-13 74-91 BL00028 16.07 9.100e-13 186-203 BL00028 16.07 8.043e-12 46-63 BL00028 16.07 8.435e-12 130-147 BL00028 16.07 9.217e-12 270-287 BL00028 16.07 6.192e-11 242-259 BL00028 16.07 4.000e-10 158-175
141	BL00501	Signal peptidases I serine proteins.	BL00501D 16.69 9.538e-14 113-133 BL00501C 9.61 8.688e-10 89-101
143	BL01020	SAR1 family proteins.	BL01020C 15.35 7.722e-20 79-130
146	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.400e-25 335-374
149	BL00126	3'5'-cyclic nucleotide phosphodiesterases proteins.	BL00126C 22.07 1.450e-25 509-550 BL00126E 35.22 3.951e-16 654-709 BL00126D 25.50 1.360e-15 565-604 BL00126B 15.20 8.200e-11 483-495 BL00126A 27.56 8.269e-11 442-479
151	BL00632	Ribosomal protein S4 proteins.	BL00632 23.79 5.271e-20 106-149
154	BL00559	Eukaryotic molybdopterin oxidoreductases proteins.	BL00559I 13.63 5.304e-19 29-58 BL00559K 13.17 2.957e-18 172-199 BL00559J 19.63 8.385e-13 99-151 BL00559L 13.60 5.814e-12 241-259
155	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 1.692e-13 13-35
157	BL00406	Actins proteins.	BL00406D 12.58 2.547e-18 275-330 BL00406A 9.95 5.776e-16 15-50 BL00406B 5.47 7.429e-12 69-124 BL00406C 6.75 9.682e-12 128-183
160	BL00132	Zinc carboxypeptidases, zinc-binding region 1 proteins.	BL00132A 26.07 7.000e-14 22-63 BL00132C 21.35 3.466e-12 104-145
165	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 9.043e-13 139-158
168	BL00362	Ribosomal protein S15 proteins.	BL00362 24.67 9.700e-15 129-172
169	BL00039	DEAD-box subfamily ATP-dependent helicases proteins.	BL00039D 21.67 1.000e-35 640-686 BL00039A 18.44 1.964e-13 212-251 BL00039B 19.19 4.553e-13 378-404 BL00039C 15.63 8.773e-12 465-489
175	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 3.721e-12 14-36
178	BL01310	ATP1G1 / PLM / MAT8 family proteins.	BL01310 14.74 2.432e-29 133-169
179	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-	PD01066 19.43 9.455e-36 6-45

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		BINDING NU.	
180	PR00007	COMPLEMENT C1Q DOMAIN SIGNATURE	PR00007B 14.16 7.429e-20 160-180 PR00007A 19.33 4.938e-19 133-160 PR00007C 15.60 1.225e-15 206-228 PR00007D 9.64 6.885e-11 238-249
181	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.526e-24 280-323
182	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.526e-24 263-306
183	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.526e-24 280-323
184	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.526e-24 263-306
188	PR00929	AT-HOOK-LIKE DOMAIN SIGNATURE	PR00929C 5.26 3.328e-09 460-471
189	PR00929	AT-HOOK-LIKE DOMAIN SIGNATURE	PR00929C 5.26 3.328e-09 440-451
190	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383F 15.51 7.188e-17 666-682 BL00383A 13.34 8.714e-17 162-177 BL00383E 10.35 1.000e-14 333-344 BL00383E 10.35 7.300e-14 628-639 BL00383F 15.51 1.720e-13 371-387 BL00383C 10.10 3.000e-13 217-228 BL00383D 11.92 7.000e-13 295-308 BL00383B 7.61 1.692e-11 187-196 BL00383C 10.10 1.750e-09 509-520 BL00383D 11.92 4.000e-09 589-602 BL00383B 7.61 8.000e-09 479-488
191	PR00450	RECOVERIN FAMILY SIGNATURE	PR00450C 12.22 7.911e-15 83-105 PR00450C 12.22 6.286e-13 47-69
193	PF00564	Octicosapeptide repeat proteins.	PF00564B 24.74 6.164e-16 227-278
194	PR00503	BROMODOMAIN SIGNATURE	PR00503D 20.81 9.156e-15 204-224 PR00503B 9.96 9.571e-13 170-187
195	BL00901	Cysteine synthase/cystathionine beta-synthase P-phosphate att.	BL00901C 20.63 3.429e-18 67-117
197	BL00636	Nt-dnaJ domain proteins.	BL00636A 8.07 6.211e-17 40-57 BL00636B 15.11 2.000e-13 67-88
198	PR00690	ADHESIN FAMILY SIGNATURE	PR00690A 10.86 9.866e-09 463-482
199	BL01131	Ribosomal RNA adenine dimethylases proteins.	BL01131A 26.62 2.343e-12 84-130
201	PR00910	LUTEOVIRUS ORF6 PROTEIN SIGNATURE	PR00910A 2.51 8.352e-12 509-522
203	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 2.286e-10 39-72
206	PR00261	LOW DENSITY LIPOPROTEIN (LDL) RECEPTOR SIGNATURE	PR00261A 11.02 4.462e-19 65-87 PR00261C 11.37 9.308e-19 65-87 PR00261D 12.47 2.667e-18 65-87 PR00261B 14.12 4.000e-18 143-165 PR00261A 11.02

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			4.833e-18 143-165 PR00261D 12.47 7.500e-18 143-165 PR00261B 14.12 5.065e-16 65-87 PR00261C 11.37 8.967e-16 143-165 PR00261F 11.57 4.938e-13 143-165 PR00261E 11.08 7.188e-13 65-87 PR00261F 11.57 7.188e-13 65-87 PR00261E 11.08 1.643e-11 143-165
209	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 6.143e-13 118-173 PF00791C 20.98 7.680e-10 132-171
211	PR00007	COMPLEMENT C1Q DOMAIN SIGNATURE	PR00007A 19.33 5.781e-19 131-158 PR00007B 14.16 4.115e-18 158-178 PR00007C 15.60 1.675e-15 201-223 PR00007D 9.64 7.231e-11 233-244
212	BL00183	Ubiquitin-conjugating enzymes proteins.	BL00183 28.97 1.545e-30 43-91
213	BL00183	Ubiquitin-conjugating enzymes proteins.	BL00183 28.97 1.545e-30 43-91
215	BL00039	DEAD-box subfamily ATP-dependent helicases proteins.	BL00039D 21.67 1.900e-29 568-614 BL00039A 18.44 1.871e-23 21-60 BL00039C 15.63 1.720e-11 364-388 BL00039B 19.19 4.064e-11 277-303
217	BL00100	Chloramphenicol acetyltransferase proteins.	BL00100D 17.22 8.484e-09 68-106
219	PR00213	MYELIN P0 PROTEIN SIGNATURE	PR00213C 15.94 3.969e-11 199-227
222	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 1.947e-09 144-155
224	PR00875	MOLLUSC METALLOTHIONEIN SIGNATURE	PR00875A 5.83 1.000e-09 901-913
225	BL00636	Nt-dnaJ domain proteins.	BL00636B 15.11 8.200e-19 18-39
226	BL00636	Nt-dnaJ domain proteins.	BL00636A 8.07 1.000e-21 21-38 BL00636B 15.11 8.200e-19 45-66
229	PR00301	70 KD HEAT SHOCK PROTEIN SIGNATURE	PR00301F 13.98 7.563e-13 329-346 PR00301G 13.78 4.300e-12 361-382
230	BL00460	Glutathione peroxidases selenocysteine proteins.	BL00460A 28.67 8.773e-20 35-70 BL00460B 9.73 7.429e-16 78-96 BL00460C 14.35 2.831e-12 111-134 BL00460D 16.89 8.773e-11 140-160
231	PR00647	SENR ORPHAN RECEPTOR SIGNATURE	PR00647B 10.19 8.522e-09 273-287
233	BL00292	Cyclins proteins.	BL00292B 20.31 7.429e-27 244-275 BL00292A 22.87 7.750e-27 201-235
234	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 6.308e-13 7-29 PR00449C